The Origin of Plants Endemic to the Mediterranean and Irano-Turanian Floristic Regions: Case Studies from the Citrus Family (Rutaceae)

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Gabriele Salvo

von

Italien

Promotionskomitee:

Prof. Dr. Elena Conti (Vorsitz)

Prof. Dr. Tony Wilson

Prof. Dr. José María Fernández-Palacios

Zürich 2010

Table of Contents

Acknowledgements

I would like to start by expressing my gratitude towards the Institute *sensu latissimo*, including its current and previous members, the librarians, the gardeners, the cleaning and cafeteria staff, and the security guards. You all created a very pleasant and comfortable atmosphere in which to work. I much enjoyed the *apéros*, the winter/summer feasts, the lunches, and the "small talk" along the building's corridors**.....** Within the Institute I would like to thank three persons in particular. First of all, Elena Conti, not just because she was my supervisor, but because she has been an excellent one: competent, stimulating, and caring. I am particularly indebted to her for the trust she reposed in me, the motivation she constantly transmitted me, and her everlasting patience and understanding. On a more practical side, I am grateful to her for having given me the opportunity and financial support to extend my Ph.D. Secondly, I would like to thank Corinne Burlet. I believe that without her this Institute would collapse! During the Ph.D. she has been so kind, helpful, and professional that no other person would do her job better than she does it. Her laughter brightened up my days and our random chitchat talking kept me always entertained. Thirdly, a big "thanks" goes to Sara Manafzadeh. She entered this Institute in a very gentle and discrete way; she has now become a sort of institution in her own right! I appreciated all the enthusiasm she transmitted and admired her patience, determination, and strength in cases where others would have lost their nerves. I feel totally indebted to her for having organized, and accompanied me in, one of the most precious journeys of my life**.....** I would like to thank all the professors and scientists I have met throughout my Ph.D., at conferences, courses, symposia, talks, herbaria, and during field work. I feel particularly grateful towards Louis Zeltner: his help in the field has been invaluable and he has been an amazing travelling companion and role model for how to make the best out of retirement! José María Fernández-Palacios and Tony Wilson have been extremely kind in being part of my Ph.D. defence committee and John Thompson and Hans-Peter Comes in reviewing my Ph.D. thesis**.....** I would like to thank the Institute, Claraz Schenkung, and SANW for their financial support for conferences and field work**.....** I would like to thank Zürich (especially the lake, the woods, Niederdorf, ASVZ, and Marché) for having provided such a relaxing, at times idyllic, setting in which to work**.....** I would like to thank all the great people I have met in Zürich for having given me the right balance outside of work**.....** I would like to thank all the good friends around the globe for their support throughout these years**.....** Finally, *dulcis in fundo*, I would like to thank my relatives and family for their constant help and inspiration.

Summary

The present thesis deals with aspects of biogeography, phylogenetics, systematics, and evolution. Its goal was to investigate patterns of biotic assembly in the Mediterranean region, with special emphasis on the effect of earth processes on species origins and distribution, the origin of species endemic to islands, and the biogeographic links between the Mediterranean and Irano-Turanian regions. To address these issues, two genera of Rutaceae (citrus family) that met several requirements were selected as model systems: *Ruta* and *Haplophyllum*. The former, the type genus of the family, is restricted to the Mediterranean region and comprises species endemic to both continental fragment (Corsica and Sardinia) and oceanic (Canary Islands) islands. The latter, one of the most species-rich genera of the family, has been used to characterize the flora of the Irano-Turanian region, where it reaches maximum species diversity, but also includes species endemic to the Mediterranean region.

In **Chapter I** we generated molecular phylogenies for *Ruta* and closely related taxa, essential to provide a robust framework for subsequent biogeographic analyses. Moreover, we tested conflicting taxonomic treatments of *Ruta* and affiliated taxa based on different classes of characters. The analyses supported the current circumscription of *Ruta* and showed that the genus can only be diagnosed by using a suite of homoplasious, plesiomorphic, and autapomorphic morphological character states. Conflict between molecular and phytochemical datasets was ascribed to convergence in secondary chemical compounds.

In **Chapter II** we carried out biogeographic analyses of *Ruta* aimed at elucidating the time frame and sequence of range expansion events associated with its origin and diversification, focusing mainly on the island endemics. Biogeographic scenarios were proposed by integrating information from phylogeny, molecular dating, and ancestral range reconstruction methods that incorporate palaeo-geographic models. The analyses showed that *Ruta* invaded the Mediterranean region from the north before the onset of the current Mediterranean climate. Land migration through a temporary connection between the Corso-Sardinian and Apulian microplates, followed by vicariance, was inferred as the process underlying the origin of the Corso-Sardinian endemic lineage. The origin and diversification of the clade restricted to the Canary Islands was explained by means of a single colonization event of the archipelago, driven by long-distance dispersal from North Africa, followed by inter-island speciation and parallel invasion of similar ecological zones.

In **Chapter III** we carried out molecular phylogenetic analyses of *Haplophyllum* in order to explore the biogeographic links between the Mediterranean and Irano-Turanian regions. The analyses identified many instances of species non-monophyly, but also cases of

strongly-supported species monophyly. Optimization of morphological characters on the molecular phylogeny indicated that several species of the genus, especially those with a widespread distribution, can only be diagnosed by combinations of homoplasious character states. Preliminary biogeographic patterns suggested that the Mediterranean representatives of the genus arrived from the east multiple times, corroborating the hypothesis that the Irano-Turanian region served as a key source for the colonization of the Mediterranean region.

Zusammenfassung

In der hier präsentierten Doktorarbeit wurden diverse Aspekte aus dem Gebiet der Biogeografie, Systematik und Evolution kombiniert, um Muster in biotischen Gemeinschaften des Mittelmeerraums zu beschreiben. Im Mittelpunkt standen hierbei geografische Prozesse, welche herangezogen wurden, um einerseits die Herkunft und Verbreitungsgebiete ausgewählter Arten und die Herkunft von Inselendemiten zu erforschen, und um andererseits zu verstehen, welche biogeografischen Elemente den Mittelmeerraum mit der irano-turanischen Region verbindet. Zur Beantwortung dieser Fragen wurden zwei Gattungen aus der Familie der Rutaceae (Zitrusgewächse) als Modellorganismen ausgewählt: *Ruta* und *Haplophyllum*. Die Gattung *Ruta*, Typusgattung der Familie, kommt nur im Mittelmeerraum vor und umfasst Endemiten von kontinentalen (Korsika und Sardinien) und vulkanischen Inseln (Kanarische Inseln). Die Gattung *Haplophyllum* ist eine der artenreichsten Gattungen der Familie. Sie charakterisiert die irano-turanische Region und umfasst zusätzlich einige Mittelmeerendemiten.

In **Kapitel 1** erstellten wir für die Gattung *Ruta* und ihre nahverwandten Taxa molekulare Phylogenien. Diese lieferten die nötigten Rahmenbedingungen für die nachfolgenden biogeografischen Analysen. Darüber hinaus konnten wir anhand unterschiedlicher Merkmals-kategorien Konflikte in der taxonomischen Zugehörigkeit von *Ruta* und vervandter Arten testen. Unsere Analysen unterstützen die derzeitige Umschreibung der Gattung *Ruta*. Zwischen dem molekularen und phytochemischen Datensatz zeigten sich Konflikte, die möglicherweise auf Konvergenzen in sekundären chemischen Stoffen zurückzuführen sind. Schliesslich zeigten unsere Analysen, dass die Gattung *Ruta* anhand einer Serie von homoplastischen, plesiomorphen und autapomorphen morphologischen Merkmalzuständen eindeutig identifiziert werden kann.

In **Kapitel 2** führten wir eine integrative biogeografische Analyse der Gattung *Ruta* durch, welche auf Phylogenien, molekularen Datierungensmethoden und Methoden zur Rekonstruktion historischer Verbreitungsgebiete mit palaeo-geografischen Modellen basierte. Diese Analyse ermöglichte Rückschlüsse auf den zeitlichen Rahmen und die Reihenfolge der einzelnen Ausdehnungsereignisse der Verbreitungsgebiete, auf die Herkunft und die Diversifikation (Artbildung) der Gattung *Ruta* und im speziellen ihrer Inselendemiten. Unsere Analysen zeigten, dass die Gattung *Ruta* vor dem Aufkommen des heutigen Mittelmeerklimas den Mittelmeerraum von Norden her besiedelte. Dies deutet darauf hin, dass die Besiedlung auf dem Landweg erfolgte, gefolgt von Vikarianz, als die Korsisch-Sardinische und die Apulische Mikroplatte kurzzeitig miteinander verbunden waren. Dadurch konnten wir die

Herkunft der Korsisch-Sardinischen Endemiten klären. Die Clades, die nur auf den Kanarischen Inseln vorkommen, besiedelten diese in einem Schritt über lange Distanzen von Nordafrika. Auf den Kanarischen Inseln selber erfolgte dann die interinsuläre Artbildung und parallele Besiedlung der Inseln entlang gleicher oder ähnlicher ökologischer Zonen.

In **Kapitel III**, untersuchten wir anhand einer molekularen Phylogenie der Gattung *Haplophyllum* wie der Mittelmeerraum mit der irano-turanischen Region biogeografisch verbunden ist. Die Analyse zeigte, dass viele Arten nicht monophyletisch, andere Arten aber deutlich monophyletisch sind. Die Optimierung von morphologischen Merkmalen auf die molekulare Phylogenie zeigte, dass viele Arten der Gattung *Haplophyllum* (vor allem diejenigen mit einem grossen Verbreitungsgebiet) nur durch die Kombination von homoplastischen Merkmalzuständen beschrieben werden können. Erste vorläufige biogeografische Muster deuten darauf hin, dass die Arten des Mittelmeerraums ihr Verbreitungsgebiet mehrfach von Osten her besiedelt haben. Dies unterstützt die Hypothese, nach welcher die irano-turanischen Region das Schlüsselgebiet war, von welchem aus die Besiedlung des Mittelmeergebietes erfolgte.

[German translation courtesy of Barbara Keller and Peter Endress]

General Introduction

The study of the geographic distribution of organisms has long fascinated biologists. In general, the restriction of a taxon to a particular geographic area is a consequence of both historical events and ecological processes (Lomolino *et al*., 2006). The former relate to the geographic origin of a species and to the spatial changes that have led to its current distribution, which is the realm of historical biogeography. The latter must be invoked to explain the present range limits of a species, which is the duty of ecological biogeography. A central task of historical biogeography is to understand how past geologic and climatic events have shaped the distribution of organisms (Crisci *et al*., 2003). The Mediterranean region, one of the 25 hotspots of biodiversity of the world (Médail and Quézel, 1997; Myers *et al*., 2000), is especially attractive to study the effect of earth processes on species origins and distribution, for its complex geologic history and palaeo-climate are well known (Dercourt *et al*., 2000; Krijgsman, 2002) and its biotic diversity is well documented (Blondel *et al*., 2010). This region, comprising the Mediterranean basin and the Macaronesian islands (Quézel, 1985), is characterized by a remarkable geological, climatological, and ecological complexity, which has been invoked to explain its impressive biodiversity (ca. 22,500 vascular plant species, of which ca. 13,000 are endemics; Thompson, 2005).

In order to understand how such biodiversity evolved, it's important to place the Mediterranean region in the context of global biogeography. Due to its location at the crossroad of Europe, Asia, and Africa, the Mediterranean region has served as the meeting ground for a complex mixture of biotic elements (Blondel *et al*., 2010). Some of these evolved *in situ*, others were filtered from the regional biota of neighbouring areas, while others arrived from more distant regions (Ackerly, 2009). The Mediterranean flora, for example, has been defined as a heterogeneous entity including elements with/of (i) tropical or subtropical affinities, (ii) northern origins (i.e., Holarctic or Eurasiatic elements), and (iii) autochtonous origins (Quézel, 1985; Thompson, 2005). Climatic changes that occurred from the Pliocene onward, including the onset of the current Mediterranean climate (characterized by dry summers and cold winters), caused the demise of most taxa belonging to the first class and the increase of plant species, mainly of a xerophytic nature, constituting the third class (Suc, 1984; Blondel *et al*., 2010). The latter include the important Irano-Turanian group, comprising taxa having centres of diversity in the semi-arid steppes of Central Asia, which penetrated the low rainfall areas of the Mediterranean region from the east during episodes of climatic change and tectonic activity (Zohary, 1973; Thompson, 2005; Blondel *et al*., 2010). In fact, based on either floristic similarities or phylogenetic evidence, several studies

suggested that the Irano-Turanian region, one of the richest areas in the Holarctic Kingdom and the richest one in South West Asia (Davis *et al*., 1994), served as a *key source* for the colonization of the Mediterranean region (Quézel, 1985, 1995; Thompson, 2005; Mansion *et al*., 2008, 2009).

Another central task of biogeography is to understand the role of islands in promoting species diversity (Wallace, 1881; MacArthur and Wilson, 1967; Whittaker and Fernandez-Palacios, 2007). Ever since Darwin and his seminal observations on the Galapagos archipelago, islands have been viewed as key laboratories of evolution (Losos and Ricklefs, 2009). Owing to their physical separation from other emerged lands, islands represent ideal systems to study the influence of barriers on speciation and patterns of colonization. The Mediterranean region, again, offers the ideal geographic setting to study biological evolution on islands, for it contains both *continental fragment islands*, derived from the splitting of moving microplates (e.g., Corsica, Sardinia, the Balearic Islands), and *oceanic islands*, formed *de novo* from the ocean floor via volcanic processes (e.g., the Canary Islands). Biotic assemblage in these two kinds of islands is fundamentally different: continental fragment islands carry with them whole communities as they split from the mainland, while oceanic islands offer a biological *tabula rasa* to the first colonizers (Whittaker and Fernandez-Palacios, 2007). Continental fragment islands have thus been viewed as museums, where remnants of ancient floristic elements, representing the autochtonous flora of the continent to which they were connected, may have survived. On the contrary, oceanic islands have been described as cradles of new biodiversity, for they present a great abundance of niches available for colonization after their emergence above the ocean (Mansion *et al*., 2009).

The aim of this thesis was to investigate patterns of biotic assembly in the Mediterranean region, with special emphasis on: i) the effect of earth processes on species origins and distribution, ii) the origin of island endemics, and iii) the biogeographic links between the Mediterranean and Irano-Turanian regions. To address the first two points we selected *Ruta* L., the type genus of Rutaceae; to investigate the last point we chose *Haplophyllum* A. Juss., one of the most species-rich, but poorly-known genera of Rutaceae. *Ruta* was selected for it is endemic to the Mediterranean region and includes species restricted to both continental fragment and oceanic islands. *R. corsica* and *R. lamarmorae* are endemic to Corsica and Sardinia, respectively, whereas *R. pinnata*, *R. microcarpa*, and *R. oreojasme* are endemic to the Canary Islands. Moreover, *Ruta* contains only nine species, allowing for complete taxon sampling, lacks traits facilitating long-distance dispersal, reducing stochastic noise, and reliable fossils for molecular clock calibration are available in Rutaceae. *Haplophyllum* was chosen for it reaches maximum species diversity in the Irano-Turanian

region, mainly in Turkey, Iran, and Central Asia, but also includes species endemic to the Mediterranean region, some as far as the Iberian peninsula. In fact, this genus was used by both Zohary (1973) and Takhtajan (1986) to characterize the flora of the Irano-Turanian region. While a detailed knowledge of the evolution of species diversity in *Haplophyllum* could yield useful insights into the biogeographic role of the Irano-Turanian region, this taxon has never been examined from a phylogenetic/biogeographic point of view. Although *Ruta* and *Haplophyllum* were selected to address the biogeographic issues mentioned above, their study opened interesting avenues of research pertaining to biological systematics, such as species circumscription, patterns of homoplasy, and character congruence.

First of all, since both the circumscription of *Ruta* and its relationships to other genera within Rutaceae have been questioned, molecular phylogenetic analyses for *Ruta* and closely related taxa (i.e., within tribe Ruteae), essential to provide a framework for subsequent biogeographic analyses, were carried out. These are presented in **Chapter I**. Additionally, since previous treatments of Ruteae based on morphology and phytochemistry contradicted each other, trees derived from morphological, phytochemical, and molecular datasets of Ruteae were compared to look for possible patterns of agreement among them. Moreover, non-molecular characters were mapped on the molecular phylogeny to identify patterns of homoplasy in the morphological and phytochemical data sets. The phylogenetic analyses supported the most recent circumscription of *Ruta*, but disagreed with more comprehensive taxonomic treatments of Ruteae. The molecular tree was congruent with the morphological one, but in conflict with the phytochemical one. Convergence in secondary chemical compounds was proposed as the possible cause of such conflict. Finally, the optimization of non-molecular characters onto the molecular tree showed that only a few morphological and phytochemical synapomorphies are congruent with the clades supported by the molecular data, while most of the morphological characters traditionally used for taxonomic purposes are homoplasious.

This study has implications on the common practice of testing traditional classifications based on morphology with new evidence gathered from molecular sequence data. Disagreement between taxonomy and phylogeny has often been attributed to high levels of homoplasy in characters traditionally used to delimit taxa and/or taxon diagnoses based on plesiomorphic morphological character states. More generally, comparisons between different sources of characters have prompted the recognition of important biological phenomena such as convergence, hybridization, and lineage sorting. Rather than being viewed in a negative sense, the identification of homoplasy in non-molecular characters should be used as a starting point to study its biological and methodological causes, focusing especially on the

developmental pathways underlying phenotypic traits and the conflicting assessments of homology often performed when comparing characters at different levels of organization.

Once a robust phylogenetic framework for *Ruta* was present, we adopted an integrative approach to historical biogeography, drawing from phylogenetics, molecular dating, and ancestral range reconstruction methods that incorporate palaeo-geographic models, in order to elucidate the time frame and sequence of range expansion events associated with the origin and diversification of *Ruta* and in particular its island endemics. This study is presented in **Chapter II**. Specifically, we asked: i) Did *Ruta* originate before the onset of the Mediterranean climate? ii) Did the tectonic splitting of the Corso-Sardinian microplate from the proto-Iberian peninsula in the Oligocene drive the origin of the *R. corsica*/*R. lamarmorae* endemic lineage? iii) Was the divergence between *R. corsica* and *R. lamarmorae* driven by the formation of the Bonifacio Strait? iv) Was the origin of *R. pinnata*, *R. oreojasme*, and *R. microcarpa* driven by the emergence of islands in the Canarian archipelago? The analyses showed that: i) *Ruta* invaded the Mediterranean region from the north before the onset of the Mediterranean climate and diversified *in situ*. ii) The origin of the *R. corsica*/*R. lamarmorae* endemic lineage is better explained by processes of land migration, followed by vicariance, through a temporary connection between the Corso-Sardinian and Apulian microplates, rather than by the splitting of the Corso-Sardinian microplate from the proto-Iberian peninsula. iii) The divergence between the two Corso-Sardinian endemics can be understood in the context of eustatic changes associated with the Messinian Salinity Crisis. iv) The monophyly of the species of *Ruta* endemic to the Canary Islands indicates a single colonization of the archipelago, driven by long-distance dispersal from North Africa, followed by inter-island speciation and asynchronous invasion of similar ecological zones.

This study demonstrates that the integration of different sources of information from phylogenetics, molecular dating, ancestral range reconstruction, and geologic/palaeo-climatic models is indispensable for explaining biogeographic patterns. Additionally, the clear formulation of a hypothesis-based framework at the onset of the research helps to avoid the construction of *a posteriori* biogeographic scenarios. With respect to island biogeography theory, this study stresses the importance of temporary land connections in the biotic assembly of continental fragment islands and of determining discrete time windows of island colonization in order to better understand distributional patterns in oceanic islands.

Finally, to explore the biogeographic links between the Mediterranean and Irano-Turanian regions, *Haplophyllum* was examined from a phylogenetic, morphological, and biogeographic standpoint. We first generated phylogenies from DNA sequences of four

regions of the chloroplast genome for 118 accessions, representing 66% of the species diversity of the genus, and then, to further improve the elucidation of species boundaries, we carried out a morphological assessment of *Haplophyllum*. These analyses are presented in **Chapter III**. The phylogenetic analyses identified many instances of species non-monophyly, but also cases of strongly-supported species monophyly. Optimization of morphological characters on the chloroplast DNA phylogeny indicated that several species of the genus, especially those with a widespread distribution, can only be diagnosed by combinations of homoplasious character states and that the main morphological characters traditionally used to classify the genus are consistent with the molecular phylogeny of *Haplophyllum*. Preliminary biogeographic patterns suggested that the Mediterranean representatives of the genus arrived from the east multiple times.

This study has allowed us to identify: i) problematic clades that require further scrutiny; ii) sister-species comparisons between widespread and narrow-endemic species, which have implications for the conservation of the latter; iii) preliminary biogeographic patterns. Additionally, the character mapping analyses have corroborated the taxonomic usefulness of the main morphological characters (i.e., number of carpels, number of ovules, and carpel dehiscence) traditionally used to classify the genus. Most importantly, this study provides a fundamental framework for more detailed phylogenetic analyses aimed at proposing sound hypotheses for the biogeographic evolution of the genus, in particular with respect to the origin of its Mediterranean representatives.

Literature cited

Ackerly, D., 2009. Some comments on the age, origin and evolution of the California and Mediterranean floras. *Journal of Biogeography* 36: 1221-1233.

Blondel, J., Aronson, J., Bodiou, J.-Y., and Boeuf, G., 2010. The Mediterranean region: biological diversity in space and time; 2nd edition. Oxford University Press, Oxford.

Crisci, J. V., Katinas, L., and Posadas, P., 2003. Historical Biogeography: an introduction. Harvard University Press, Cambridge, MA.

Davis, S. D., Heywood, V. H., and Hamilton, A. C., 1994. Centres of plant diversity: a guide and strategy for their conservation. 3 Vol. IUCN Publications Unit, Cambridge.

Dercourt, J., Gaetani, M., Vrielynck, B., Barrier, E., Biju-Duval, B., Brunet, M. F., Cadet, J. P., Crasquin, S., and Sandulescu, M. e., 2000. Atlas Peri-Tethys, Palaeogeographical maps. CCGM/CGMW, Paris.

Krijgsman, W., 2002. The Mediterranean: *mare nostrum* of Earth sciences. *Earth and Planetary Science Letters* 205: 1-12.

Lomolino, M. V., Riddle, B. R., and Brown, J. H., 2006. Biogeography; 3rd edition. Sinauer Associates, Sunderland, MA.

Losos, J. B. and Ricklefs, R. E., 2009. Adaptation and diversification on islands. *Nature* 457: 830-836.

MacArthur, R. H. and Wilson, E. O., 1967. The theory of island biogeography. Princeton University Press, Princeton, NJ.

Mansion, G., Rosenbaum, G., Schoenenberger, J., Bacchetta, G., Rossello, J., and Conti, E., 2008. Invasions of the Mediterranean basin by the Araceae: integrative phylogeny-based evidence. *Systematic Biology* 57: 269-285.

Mansion, G., Selvi, F., Guggisberg, A., and Conti, E., 2009. Origin of Mediterranean insular endemics in the Boraginales: integrative evidence from molecular dating and ancestral area reconstruction. *Journal of Biogeography* 36: 1282-1296.

Médail, F. and Quézel, P., 1997. Hot-spots analysis for conservation of plant biodiversity in the Mediterranean basin. *Annals of the Missouri Botanical Garden* 84: 112-127.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., and Kent, J., 2000. Biodiversity hotspots for conservation priorities. *Nature* 403: 853-858.

Quézel, P., 1985. Definition of the Mediterranean region and the origin of its flora. Pp. 9-24 in: Gómez-Campo, C. (Ed.), Plant conservation in the Mediterranean area. Dr. W. Junk Publishers, Dordrecht.

Quézel, P., 1995. La flore du bassin méditerranéen: origine, mise en place, endémisme. *Ecol. Medit.* 21: 19-39.

Suc, J.-P., 1984. Origin and evolution of the Mediterranean vegetation and climate in Europe. *Nature* 307: 429-432.

Takhtajan, A., 1986. Floristic regions of the world. University of California Press, Berkeley.

Thompson, J. D., 2005. Plant evolution in the Mediterranean. Oxford University Press Inc., New York.

Wallace, A. R., 1881. Island life. Harper & Bros., New York.

Whittaker, R. J. and Fernandez-Palacios, J. M., 2007. Island biogeography: ecology, evolution, and conservation; 2nd edition. Oxford University Press Inc., New York.

Zohary, M., 1973. Geobotanical foundations of the Middle East. Vol. I and II. Gustav Fischer Verlag, Stuttgart; Swets & Zeitlinger, Amsterdam.

Chapter I

Phylogenetic relationships of Ruteae (Rutaceae): new evidence from the chloroplast genome and comparisons with non-molecular data

Gabriele Salvo[*1], Gianluigi Bacchetta[2], Farrokh Ghahremaninejad[3], Elena Conti[1]

[1]*Institute of Systematic Botany, University of Zürich, Zollikerstrasse 107, CH-8008 Zürich, Switzerland.*

[2]*Center for Conservation of Biodiversity (CCB) and Department of Botany, University of Cagliari, Viale S. Ignazio da Laconi 13, 09123 Cagliari, Italy.*

[3]*Department of Biology, Tarbiat Moallem University, 49 Dr. Mofatteh Avenue, 15614 Tehran, Iran.*

Published in:

Molecular Phylogenetics and Evolution 49: 736-748 (2008)

Abstract

Phylogenetic analyses of three cpDNA markers (*mat*K, *rpl*16, and *trn*L-*trn*F) were performed to evaluate previous treatments of Ruteae based on morphology and phytochemistry that contradicted each other, especially regarding the taxonomic status of *Haplophyllum* and *Dictamnus*. Trees derived from morphological, phytochemical, and molecular datasets of Ruteae were then compared to look for possible patterns of agreement among them. Furthermore, non-molecular characters were mapped on the molecular phylogeny to identify uniquely derived states and patterns of homoplasy in the morphological and phytochemical data sets. The phylogenetic analyses determined that *Haplophyllum* and *Ruta* form reciprocally exclusive monophyletic groups and that *Dictamnus* is not closely related to the other genera of Ruteae. The different types of data sets were partly incongruent with each other. The discordant phylogenetic patterns between the phytochemical and molecular trees might be best explained in terms of convergence in secondary chemical compounds. Finally, only a few non-molecular synapomorphies provided support for the clades of the molecular tree, while most of the morphological characters traditionally used for taxonomic purposes were found to be homoplasious. Within the context of the phylogenetic relationships supported by molecular data, *Ruta*, the type genus for the family, can only be diagnosed by using a combination of plesiomorphic, homoplasious, and autapomorphic morphological character states.

Keywords

Ruta, citrus family, morphology, phytochemistry, congruence, Shimodaira-Hasegawa test, character mapping, homoplasy.

Introduction

Testing whether traditional taxonomic classifications based on morphology are congruent with more recent molecular phylogenetic findings has become a central task in the current systematic agenda (e.g., Simões *et al*., 2004; Van der Niet *et al*., 2005; Wiens *et al*., 2005; Marazzi *et al*., 2006; Rønsted *et al*., 2007; but see Grant, 2003). Disagreements between morphological taxonomies and molecular phylogenies have often been attributed to high levels of homoplasy in characters traditionally used to delimit taxa (e.g., Lavin *et al*., 2001; Moylan *et al*., 2004; Mueller *et al*., 2004; Simões *et al*., 2006) and taxon diagnoses based on plesiomorphic morphological character-states (e.g., Roalson *et al*., 2005; Norup *et al*., 2006). Incongruence between molecular phylogenies and morphological classifications has prompted the recognition of groups highly supported by molecular data, but lacking unique morphological synapomorphies (e.g., Porter and Johnson, 2000; Lavin *et al*., 2001; Hughes *et al*., 2004), or the dismantling of traditionally accepted taxa (e.g., Kim *et al*., 1996; Kron *et al*., 1999; Wiens *et al*., 2005).

More generally, the choice of characters for phylogenetic analysis has been a crucial and controversial issue in systematics (e.g., Hart *et al*., 2004; Stace, 2005) and the relative role of molecular and morphological data in reconstructing phylogenies has been extensively debated (Hillis, 1987; Patterson, 1988; Sytsma, 1990; Donoghue and Sanderson, 1992; Novacek, 1994; Baker *et al*., 1998; Wahlberg and Nylin, 2003; Wortley and Scotland, 2006). Directly linked to character choice is the controversy about combined versus separate analyses of different datasets (Bull *et al*., 1993; de Queiroz *et al*., 1995). For example, should morphological, molecular, and phytochemical characters for a certain group of organisms be analyzed together or separately? Advocates of separate analyses have stressed the fact that congruence among trees derived from independent sources of data can offer strong evidence for the accuracy of the inferred relationships (Swofford, 1991; Hillis, 1995; Miyamoto and Fitch, 1995; Graham *et al*., 1998), while incongruence can provide initial insights on important biological phenomena, ranging from hybridization to lineage sorting (e.g., Rieseberg *et al*., 1996; Won and Renner, 2003; Doyle *et al*., 2004). Conversely, advocates of global evidence have emphasized the fact that combining datasets before phylogenetic analysis grants the best opportunity to resolve relationships at different scales of divergence (Cunningham, 1997; Kluge, 1998; Gatesy and Baker, 2005).

Ruta L. (Rutaceae Juss.) and related genera offer a primary example of the discordant systematic conclusions that can be reached by using different types of data. Below we provide the necessary background to understand the sources of such discrepancy and explain how novel evidence from molecular characters might help to clarify the discordant taxonomic

treatments published until now. The paucity of diagnostic morphological traits, combined with their overlapping and contradicting nature, has hindered both a stable circumscription for *Ruta* – alternately subjected to taxonomic "lumping" (Engler, 1896, 1931) and "splitting" (Townsend, 1968, 1986) – and the unequivocal identification of relationships with other genera of Rutaceae (Townsend, 1986).

At the family level, Rutaceae (161 genera/1815 species; Stevens, P.F., 2001 onwards, Angiosperm Phylogeny Website) have been investigated morphologically (Engler, 1896, 1931; Saunders, 1934; Moore, 1936; Scholz, 1964; Tilak and Nene, 1978), molecularly (Chase *et al*., 1999; Scott *et al*., 2000; Samuel et *al*., 2001; Morton *et al*., 2003), and biochemically, owing to their remarkable diversity of secondary chemical compounds (Price, 1963; Fish and Waterman, 1973; Waterman 1975, 1983, 1990; Gray and Waterman, 1978; Waterman and Grundon, 1983; Kong *et al*., 1986; Ng *et al*., 1987; Da Silva *et al*., 1988; Zakaria, 2001). However, different types of characters led to contrasting systematic conclusions. For example, some taxonomic groups recognized in the most comprehensive morphological study (Engler, 1896, 1931) and the most recent chemotaxonomic survey (Da Silva *et al*., 1988) of Rutaceae conflict with each other and with the groups supported in the broadest molecular phylogenies available until now (Chase *et al*., 1999; Scott *et al*., 2000). The cited molecular studies, based on sparse character and taxon sampling, supported *Ruta* either as sister to a genus of subfamily Flindersioideae (Chase *et al*., 1999), or as sister to a clade of subfamily Citroideae (Scott *et al*., 2000), while Engler (1896, 1931) had placed it within subfamily Rutoideae.

In his comprehensive morphological study of Rutaceae, Engler (1896, 1931) divided tribe Ruteae into two subtribes: Rutinae, comprising *Ruta*, *Thamnosma* Torrey and Frémont, *Boenninghausenia* Reichb. ex Meissner, *Cneoridium* Hook.f., and *Psilopeganum* Hemsl. ex Forb. and Hemsl.; and Dictamninae, consisting only of *Dictamnus* L. (Table 1). Furthermore, he split *Ruta* into subgenus *Euruta* Engl., housing five species, three of which were originally described by Linnaeus (1735; 1753), and subgenus *Haplophyllum* (around 50 species; see Table 1). Later systematic treatments (Mester and Vicol, 1971; Townsend, 1986; Da Silva *et al*., 1988; Navarro *et al*., 2004), however, ranked *Haplophyllum* at the generic level, as originally proposed by Jussieu (1825), reducing the number of species in *Ruta* from around 60 to eight, as currently recognized (Townsend, 1968; Bramwell and Bramwell, 2001).

The six genera included in Ruteae by Engler (1896, 1931) were each distinguished by the following morphological traits (Table 1): *Ruta* (around 60 species) by tetra- and pentamerous flowers, a thick cushion-shaped nectary disk, and dorsally angled seeds; *Thamnosma* (one species) by almost reniform seeds and variation in the shape of the nectary disk;

Boenninghausenia (one species) by a cup-shaped nectary disk and filiform filaments; *Cneoridium* (one species) by one carpel, two ovules per locule, and an almost spherical stigma; *Psilopeganum* (one species) by a relatively small nectary disk with a narrow ending; and *Dictamnus* (one species) by zygomorphic flowers, lanceolate petals and sepals, club-shaped filaments with protruding glands, and three ovules per locule (see Table 2). *Psilopeganum* was analyzed in a systematic context only by Engler (1836, 1931), but its narrow occurrence in the Three Gorges Reservoir area of central China (Song *et al*., 2004; Tang *et al*., 2007) prevented its inclusion in more recent taxonomic treatments (e.g., Townsend, 1986; Da Silva *et al*., 1988).

Despite the systematic importance of the above-mentioned diagnostic features, relationships and taxonomic boundaries among the six genera of Ruteae (Engler, 1896, 1931) remain controversial. Townsend (1986) observed that the states of some characters traditionally used to differentiate the genera overlap or suggest contradicting sister-group relationships (Table 2). For example, the ranges of the number of ovules per locule overlap across *Ruta* and allied genera. The presence of cuneate filaments favours *Cneoridium* and *Thamnosma* as sister taxa, whereas spherical seeds link *Cneoridium* with *Dictamnus*. Moreover, Townsend (1986) argued that there are no grounds for considering *Haplophyllum* to be more closely related to *Ruta* than to *Thamnosma*, as proposed by Engler (1896, 1931). In fact, while *Ruta* and *Haplophyllum* share several morphological similarities, including translucent dots on the leaves, yellowish flowers, a thick nectary disk, a short thick style, and connate carpels, they can be clearly distinguished by differences in petal margins, flower merism, seed shape, and pollen morphology. Furthermore, Townsend (1986) showed that the pollen grains of *Ruta* and *Thamnosma* are more morphologically similar to each other than to those of *Haplophyllum*.

The inclusion of *Dictamnus albus* L., the only species of the genus *Dictamnus* and subtribe Dictamninae, in Ruteae (Engler 1896, 1931; Table 1) is also contentious, for this species is distinct from all other Rutaceae due to the presence of special quinolones and limonoids and the absence of coumarins (Da Silva *et al*., 1988). Furthermore, Moore (1936) remarked that the floral anatomical differences between *Dictamnus* and *Ruta* are greater than those between any two genera within any other tribe of Rutaceae, thus criticizing the inclusion of *Dictamnus* and *Ruta* in Ruteae. Therefore, considering the above-mentioned criticisms towards Engler's (1896, 1931) classification of Ruteae, Townsend (1986) called for a comprehensive systematic re-examination of the entire tribe.

Among the genera of Ruteae (Engler, 1896, 1931), *Ruta* is characterized by strong-smelling ethereal oils in its leaves, greenish-yellow petals with dentate or fimbriate margins,

and inflorescences with pentamerous terminal flowers and tetramerous lateral flowers (Townsend, 1968). As currently circumscribed (Townsend, 1968; Bramwell and Bramwell, 2001), *Ruta* includes eight species of perennial shrubs, with four species widely distributed in the Mediterranean (*R. chalepensis* L., *R. graveolens* L., *R. angustifolia* Pers., *R. montana* (L.) L.), one species endemic to the islands of Corsica and Sardinia (*R. corsica* DC.), and three species endemic to the Canary Islands (*R. pinnata* L.f., *R. oreojasme* Webb and Berth., *R. microcarpa* Svent.). Recently, the populations of *R. corsica* from Sardinia have been described as a ninth species, *R. lamarmorae*, based on morphological, karyological, and ecological differences with the populations of *R. corsica* from Corsica (Bacchetta *et al*., 2006).

Overall, morphological data have not been successful in elucidating the relationships and taxonomic boundaries of Ruteae owing to (i) the paucity of characters diagnostic for the genera within Ruteae, (ii) the conflicting and overlapping nature of the characters traditionally used to establish relationships within Ruteae, and (iii) the different taxonomic value assigned by different authors to comparative characters. With respect to Engler's (1896, 1931) classification of Ruteae, the most controversial issues are the placement of *Haplophyllum* within *Ruta* and the inclusion of *Dictamnus* in Ruteae (Townsend, 1986; Da Silva *et al*., 1988; see Table 1).

Phytochemical characters have also been used to generate taxonomic treatments of Rutaceae (e.g., Kong *et al*., 1986; Ng *et al*., 1987; Samuel *et al*., 2001; Zakaria, 2001), even though they pose specific problems that appear to limit their taxonomic value. Firstly, phytochemical information on the family is fragmentary, with only 30% of the species of Rutaceae examined (Waterman, 1990). Secondly, convergence in the production of secondary chemical compounds has been regarded as a primary source of erroneous systematic conclusions (Hegnauer, 1966; Mothes, 1981; Waterman, 1990; Waterman, 1998). For example, Waterman and Grundon (1983) argued that the synthesis of carbazole, benzophenanthridine, and quinolone alkaloids, occurring in taxa of Rutaceae with little or no immediate affinity with each other, originated by convergent evolution.

Considering the controversial interpretation of morphological (Townsend, 1986) and biochemical characters (Da Silva *et al*., 1988; Waterman, 1990), which produced contradictory taxonomic treatments for Ruteae (Table 1), and the conflict between traditional taxonomies (Engler, 1896, 1931) and available molecular phylogenies of Rutaceae (Chase *et al*., 1999; Scott *et al*., 2000), we performed a detailed phylogenetic study based on sequences from three chloroplast DNA markers to address the following questions: 1) Does the molecular phylogeny support Engler's (1896, 1931) circumscription of Ruteae and,

specifically, the inclusion of *Dictamnus* in the tribe? 2) Does the molecular phylogeny support Engler's (1896, 1931) circumscription of *Ruta* and, specifically, the treatment of *Haplophyllum* as a subgenus of *Ruta*? 3) Does the molecular phylogeny support the newly described species *R. lamarmorae*? 4) Do phylogenetic analyses of morphological, phytochemical, and DNA sequence data yield the same or different relationships among *Ruta* and closely related genera? 5) Which morphological and/or phytochemical characters are congruent with the clades recovered from the molecular phylogenetic analysis? More generally, the discussion of our results on the phylogeny of Ruteae provides an opportunity to elaborate on the sources of discrepancy among different types of data, one of the fundamental debates in systematics.

Material and Methods

Taxon sampling

Ruta and its most closely related genera (Engler, 1896, 1931; Townsend, 1986; Da Silva *et al*., 1988), with the exception of *Psilopeganum* (1 sp.), were sampled: *Ruta* (8/8 species), *Haplophyllum* (24/66 species), *Thamnosma* (5/9 species), *Boenninghausenia* (1/1 sp.), *Cneoridium* (1/1 sp.), and *Dictamnus* (1/1 sp.; see Appendix 1). It was impossible to sample *Psilopeganum sinense*, because it is restricted to the Three Gorges Reservoir area, in the Hubei province of central China, and is endangered (Song *et al*., 2004; Tang *et al*., 2007). Furthermore, this taxon is poorly represented in the herbaria that were visited during the duration of the present study (i.e., W, LE, P, BR). In order to elucidate relationships within *Ruta*, different accessions from the eight species of the genus were selected. Five accessions of *R. corsica* from Corsica and Sardinia were sampled to verify the treatment of the populations from Sardinia as a separate species (i.e., *R. lamarmorae*; Bacchetta *et al*., 2006). To test the monophyly of Ruteae (Engler, 1896, 1931), in particular the inclusion of *Dictamnus* in the tribe, eight taxa outside the tribe, belonging to subfamilies Rutoideae and Toddalioideae, were selected. Choice of outgroups was guided by Engler's (1896, 1931) classification of Rutaceae and previous phylogenetic findings (Chase *et al*., 1999; Scott *et al*., 2000). Because the molecular phylogenetic analysis of Scott *et al*. (2000) placed *Ruta* as sister to members of subfamily Aurantioideae, rather than Rutoideae, as suggested by Engler (1896, 1931), and because the monophyly of the tribes and subfamilies of Rutaceae proposed by Engler (1896, 1931) has been questioned (Da Silva *et al*., 1988; Chase *et al*., 1999), taxa from Meliaceae and Simaroubaceae, closely related to Rutaceae (Gadek *et al*., 1996; Muellner *et al*., in press), were chosen as outgroups, to reduce the possibility that the rooting taxa might fall within the ingroup. The final matrix contained 73 accessions: 66 belonging to Rutaceae,

four to Meliaceae, and three to Simaroubaceae. Included material, voucher information, sources, and GenBank/EBI accession numbers are listed in Appendix 1.

Character sampling

After performing preliminary analyses with different cpDNA markers, three markers that provided sufficient resolution at our level of investigation and allowed unequivocal alignments were chosen: the *mat*K gene, the *rpl*16 intron, and the *trn*L-*trn*F intergenic spacer, which already proved effective at resolving inter-generic relationships in other groups of angiosperms (Simões *et al*., 2004; Guggisberg *et al*., 2006; Marazzi *et al*., 2006; Rutschmann *et al*., 2007).

DNA extraction, amplification and sequencing

Prior to DNA extraction, silica-dried leaf material (15-20 mg) was ground using glass beads and a MM 3000 shaker (Retsch GmbH, Haan, Germany). Total genomic DNA was extracted using the DNeasy Plant Mini Kit from QIAGEN AG (Basel, Switzerland), following the manufacturer's instructions. The *mat*K cpDNA coding region was amplified using primers 1F and 1R (Sang *et al*., 1997). The *rpl*16 intron was amplified using primers F71 and R1516 (Baum *et al*., 1998). The *trn*L-*trn*F spacer was amplified with the primers e and f (Taberlet *et al*., 1991). All PCR reactions were 20 µl in volume. Each reaction included 9.2 µl of ddH_2O, 2 µl of Taq-Buffer [10x, 15mM $MgCl_2$], 1.6 µl of $MgCl_2$ [25 mM], 3.2 µl of dNTP [1.25 mM], 0.2 µl of Taq-Polymerase [5U/µl], 1 µl of BSA, 0.4 µl of each primer (forward and reverse), and 2 µl of DNA template. Amplification of the *mat*K region consisted of 2 min at 94°C followed by 30 cycles of: 1.5 min denaturation (94°C), 2 min annealing (53°C), and 3 min extension (72°C). After the last cycle, the temperature was kept at 72°C for the last 15 min of extension and then lowered to 4°C. Amplification of both the *rpl*16 and *trn*L-*trn*F regions consisted of 2 min at 94°C followed by 35 cycles of: 0.5 min denaturation (94°C), 1 min annealing (52°C), and 1.75 min extension (72°C). After the last cycle the temperature was kept at 72°C for 10 min of extension and then lowered to 4°C. All PCR and cycle sequencing reactions were run on a TGradient thermocycler (Biometra, Goettingen, Germany). In order to detect amplified DNA target regions and possible contamination, PCR products were separated on 1% agarose gels, stained with ethidium bromide, and viewed under UV light. Successfully amplified products were purified with the GFX PCR DNA and Gel Band purification Kit (Bioscience Amersham, Otelfingen, Switzerland), following the manufacturer's recommendations.

Cycle sequencing reactions were carried out using the BigDye™ Terminator Mix (Applied Biosystems, Inc., Foster City, California) and the same primers as above. The sequencing protocol consisted of 24 cycles of 10 sec denaturation (96°C), 5 sec annealing

(50°C), and 4 min elongation (60°C). Products were run on an ABI 3100 Genetic Analyzer (Applied Biosystems, Inc., Foster City, California) according to the manufacturer's protocol. For each region both strands were sequenced.

Alignment and phylogenetic analyses

Sequences were edited and assembled using Sequencher 4.2™ software (Gene Codes Corp., Ann Arbor, Michigan, USA). Base positions were individually double-checked for agreement between the complementary strands. All sequences were visually aligned in MacClade 4.06 (Maddison and Maddison, 2000). Regions of ambiguous alignment were excluded from the analysis (Kelchner, 2000). Gap positions were treated as missing data, unequivocally aligned gaps being coded as presence/absence of characters with the software GapCoder (Young and Healy, 2003) and then added as binary characters to the data matrix.

Three data partitions were defined, corresponding to the three loci of the chloroplast genome examined in this study. The individual partitions were initially analyzed separately to establish whether there were any strongly supported (i.e., > 85 bootstrap percentage, BP), incongruent clades among the respective trees. Since no such incongruence was detected (see Results), the sequences of the three loci were combined in a single dataset. The combined matrix was then analyzed phylogenetically either with the gaps treated as missing data ("combined without gap coding") or with the gaps coded in GapCoder (Young and Healy, 2003; "combined with gap coding").

The individual partitions and the combined matrix without gap coding were analyzed using maximum parsimony (MP). The combined matrix with gap coding was analyzed using both MP and Bayesian Inference (BI). Parsimony analyses were conducted using PAUP*4.0b10 (Swofford, 2001). All changes were treated as unordered and equally weighted (Fitch, 1971). Tree search was performed using the following protocol: (i) a heuristic search was carried out with 1000 replicates of random taxon addition sequence and 10 trees held at each step, and tree bisection-reconnection branch swapping (TBR) on best trees only, with no more than 100 trees saved per replicate; (ii) the best trees found in (i) were then used as starting trees for a second heuristic search using TBR branch swapping until all swapping options were explored, and saving multiple trees (MULTREES option in effect). The STEEPEST DESCENT option was used in both (i) and (ii). Relative levels of homoplasy in all partitions were assessed using the consistency index (CI) and the retention index (RI) as implemented in PAUP*4.0b10 (Swofford, 2001).

Relative support for each node obtained by MP was assessed using bootstrap re-sampling (Felsenstein, 1985). The following protocol was employed: heuristic search, 1000 bootstrap replicates, 100 random addition sequence replicates with three trees held at each

step, TBR swapping with STEEPEST DESCENT and saving no more than ten trees per replicate.

Bayesian inference was performed in MRBAYES v3.1.2 (Hulsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003), after determining the model of evolution most suitable for each individual cpDNA region with the Akaike Information Criterion (AIC; Akaike, 1974) in ModelTest 3.06 (Posada and Crandall, 1998). Subsequently, the commands "lset NST=6, RATES=gamma" and "lset coding=variable" were entered in MRBAYES v3.1.2 for the nucleotide and gap characters, respectively (Ronquist and Huelsenbeck, 2003). The analysis was performed with four Monte Carlo Markov chains (one cold and three incrementally heated) run for 5 x 10^6 generations, with trees sampled every 1000^{th} generation (NGEN=1.000.000, PRINTFREQ=1000, SAMPLEFREQ=1000, NCHAINS=4). Each chain used a random tree as starting point and the default temperature parameter value of 0.2. Two independent Markov chain Monte Carlo runs were carried out to check for convergence on the same region of tree space. The first 25000 sampled trees were discarded as "burn in" after checking for stability on the log-likelihood curves. The remaining trees were imported into PAUP*4.0b10 (Swofford, 2001) and used to build a 95% majority rule consensus tree showing the posterior probabilities (PP) of all observed bi-partitions.

Analyses of constrained topologies

Engler's (1896, 1931) classification of tribe Ruteae has been criticized with respect to two points: the placement of *Dictamnus* in the tribe (Moore, 1936; Da Silva *et al*., 1988) and the treatment of *Haplophyllum* as a subgenus of *Ruta* (Townsend, 1986). Two topological constraints were thus defined: (i) *Dictamnus* within Ruteae and (ii) *Haplophyllum* as sister to *Ruta*. For each constraint, a heuristic search was performed to find the shortest trees, which were then assigned a likelihood score (Table 3) using the parameters previously estimated (see Table 4). The likelihood scores of the 94 MP trees with constraint (i) and of the 940 MP trees with constraint (ii) were each compared with the highest likelihood score of the MP unconstrained trees, for a total of 1034 pair-wise comparisons, and the differences between likelihood scores calculated (Shimodaira-Hasegawa (S-H) test; Shimodaira and Hasegawa, 1999; Goldman *et al*., 2000; Lee and Hugall, 2003). Significance levels for the differences were checked on a distribution generated by using a RELL technique (resampling estimated log likelihoods; Kishino and Hasegawa, 1989). The S-H tests were performed in PAUP*4.0b10 (Swofford, 2001) on the matrix without gap coding, because this software does not allow for likelihood estimation in datasets containing two or more partitions explained by different models of evolution.

Phylogenetic comparisons among different datasets

To evaluate whether different types of data produce congruent phylogenies, we generated comparable matrices from morphology, biochemistry, and DNA sequences. Published information was available mainly at the genus level for the morphology of Rutaceae (Engler, 1896, 1931; Saunders, 1934; Scholz, 1964), and at the species level for the secondary chemical compounds (e.g., Price, 1963; Fish and Waterman, 1973; Waterman 1975, 1983, 1990; Gray and Waterman, 1978; Waterman and Grundon, 1983; Kong *et al*., 1986; Da Silva *et al*., 1988), but not for all the species included in our molecular analyses. Therefore, comparisons among different datasets were performed at the genus level and only on *Haplophyllum*, *Thamnosma*, *Boenninghausenia*, *Cneoridium*, and *Ruta*, because the results of our molecular analyses excluded *Dictamnus* from Ruteae, in agreement with Moore (1936) and Da Silva *et al*. (1988; see Results below). The genus *Choisya*, belonging to the subfamily Rutoideae (Engler, 1896, 1931), was chosen to root the resulting trees, because it is closely related to Ruteae (Da Silva *et al*., 1988) and both morphological and phytochemical data were available for it.

Morphological characters that vary among the six genera listed above were selected from descriptions in Engler (1931) and Townsend (1986), scored for different states (i.e., multistate), and used to build a matrix (see Appendix 2). The number of ovules per locule was not included in the matrix of Appendix 2, because states overlapped extensively. Phytochemical characters were selected from Da Silva *et al*. (1988), the most recent and comprehensive chemotaxonomic survey of Rutaceae. These characters referred to the presence or absence of specific phytochemical compounds and consequently were scored as binary data (Appendix 3). The molecular dataset was built by keeping only one, randomly chosen exemplar sequence from the combined matrix with coded gaps for each of the six genera. The selected sequences (indicated by an asterix in Appendix 1) were visually re-aligned in MacClade 4.06 (Maddison and Maddison, 2000). A global dataset was also constructed by combining all types of characters. For each of the four datasets, exhaustive searches were carried out using maximum parsimony with the program PAUP*4.0b10 (Swofford, 2001). Branch support was calculated with 10,000 bootstrap replicates using a Branch and Bound search strategy and an “as is” taxon addition sequence.

The trees resulting from the molecular, morphological, and phytochemical datasets were compared with each other in two ways. First, because phylogenetic incongruence should only include cases where conflicting clades are strongly supported, topologies were evaluated by direct node-to-node comparisons using branch support values (e.g., Mason-Gamer and Kellogg, 1996; Graham *et al*., 1998). Node-to-node comparisons were executed by using the

table of "bipartitions found in one or more trees and frequency of occurrence" from the bootstrap output produced in PAUP*4.0b10 (Swofford, 2001). Secondly, the trees were compared by means of the incongruence length difference (ILD) test (Farris *et al*., 1994), implemented as the "partition homogeneity" test in PAUP*4.0b10 (Swofford, 2001). For the ILD test, 1000 random repartitions were used and a branch and bound tree search was implemented. Following suggestions by Cunningham (1997) and Lee (2001), the test was carried out after removing uninformative characters.

Character mapping

We investigated whether any of the morphological and/or phytochemical characters listed in Appendices 2 and 3 were congruent with the relationships inferred from the analysis of the 6-taxon molecular dataset. Non-molecular characters were mapped on the molecular topology using maximum parsimony and both accelerated (ACCTRAN) and delayed (DELTRAN) character state optimizations, as implemented in MacClade 4.06 (Maddison and Maddison, 2000).

Results

Alignment and phylogenetic analyses

Aligned lengths, character and tree statistics, CI and RI values for all five partitions are summarized in Table 4. Forty-six and 27 nucleotide positions were excluded from the *rpl*16 and *trn*L-*trn*F partitions, respectively, owing to ambiguities in the alignment caused by strings of mononucleotides (Kelchner, 2000). Among the three cpDNA partitions, the *mat*K dataset contained the highest proportion of parsimony-informative characters (27.4%) and produced the best resolved tree, whereas the *trn*L-*trn*F partition had the highest CI and RI values (Table 4). Because no strongly supported (> 85 BP) incongruent clades were detected among individual trees, the three partitions were combined, producing an alignment of 3594 characters. The gaps of the combined matrix yielded 204 additional characters, for a total of 3798 characters (Table 4). The combined matrix with gaps coded was used for final MP and Bayesian analyses, because it had a higher number of parsimony informative sites and similar CI and RI values as compared with the combined matrix without coded gaps (see Table 4).

For all three DNA regions the AIC of ModelTest selected the same model of evolution, TVM + G, which was implemented in the Bayesian analysis, with parameters estimated again from the data. The two runs of the Bayesian analysis produced identical 50% majority-rule consensus trees, suggesting convergence on the same region of tree space. Each run reached a stationary likelihood after approximately 280.000 generations, which were not used to build the Bayesian consensus tree. The 95% majority-rule consensus tree obtained

from one run of the Bayesian analysis, with posterior probabilities (PP), and bootstrap percentages (BP) obtained by bootstrapping the same matrix under parsimony, is shown in Figure 1. This tree was similar to the strict consensus tree found from the MP search of the same matrix (not shown), except that in the MP tree some groups within *Haplophyllum* were resolved differently compared to the Bayesian tree.

The main phylogenetic results are the following (Fig. 1): 1) *Ruta*, *Boenninghausenia*, *Thamnosma*, *Haplophyllum*, and *Cneoridium* form a monophyletic group with maximum support (100 BP and 1.00 PP); 2) *Dictamnus* forms a clade with *Skimmia* Thunb. and *Orixa* Thunb. (100 BP and 1.00 PP); 3) the eight species currently ascribed to *Ruta* (Townsend, 1968; Bramwell and Bramwell, 2001) form a strongly supported clade (100 BP and 1.00 PP) that is sister to a clade consisting of *Thamnosma* and *Boenninghausenia* (100 BP and 1.00 PP), while *Haplophyllum* is sister to *Cneoridium* (100 BP and 1.00 PP); 4) *R. chalepensis* and *R. angustifolia* form a clade sister to *R. graveolens* and these three species are sister to *R. corsica* (clade I); 5) the three *Ruta* species from the Canary Islands (*R. pinnata*, *R. microcarpa*, and *R. oreojasme*) form a strongly supported clade (100 BP and 1.00 PP; clade III); 6) the relationship between *R. montana* (clade II; Fig. 1) and clades I and III described above remains unresolved; 7) within *R. corsica* there is a strongly supported split between the samples from Sardinia (86 BP and 1.00 PP) and those from Corsica (99 BP and 1.00 PP).

Analyses on constrained topologies

When *Dictamnus* was forced inside Ruteae (constraint i) and when *Haplophyllum* was constrained to be sister to *Ruta* (constraint ii), 94 MP trees of 2388 steps and 940 MP trees of 2455 steps were found, respectively, as compared to the 6 MP trees of 2092 steps resulting from the unconstrained search (Table 3). All pair-wise comparisons involving constraint (ii) produced significant differences between likelihood scores ($P < 0.001$; Table 3, in bold). However, none of the comparisons involving constraint (i) produced significant results (P values between 0.080 and 0.038; Table 3).

Comparisons among different datasets

For each six-taxon dataset, only one MP tree was found (Fig. 2; Table 5). The inter-generic relationships inferred from the six-taxon molecular dataset (Fig. 2a) were identical to those generated from the 73-accession matrix (Fig. 1), with *Ruta* sister to *Boenninghausenia*/*Thamnosma* and these three genera, in turn, sister to *Haplophyllum*/*Cneoridium*, indicating that random selection of one exemplar per genus did not change the phylogenetic pattern. All the nodes of the molecular tree were highly supported (Fig. 2a), whereas only one node each received high BP value in the morphological (Fig. 2b) and molecular trees (Fig. 2c). The relationships of the global tree were identical to

those of the molecular tree and strongly supported (Fig. 2d). The *Boenninghausenia*/*Thamnosma* clade received maximum support (100 BP) from the molecular data and was found in 86% of the bootstrap replicates of the morphological data. The *Haplophyllum*/*Cneoridium* clade, with 99 BP in the molecular tree, was found in none of the bootstrap replicates of the non-molecular datasets. The phytochemical tree shared no clades with the molecular tree and its single strongly supported clade (i.e., *Ruta*/*Haplophyllum*; 97 BP) received low bootstrap support (66 BP) in the morphological tree.

Based on the results of the ILD test, the molecular dataset was found to be significantly incongruent with both the morphological (P = 0.002) and phytochemical (P = 0.0001) datasets, while the morphological and phytochemical datasets were not significantly incongruent with each other (P = 0.41).

Character mapping

Sixteen of the 47 non-molecular characters mapped on the molecular tree exhibited equivocal reconstructions, that is, character-state transitions occurred in different branches depending on whether ACCTRAN or DELTRAN were used for the optimization. The remaining 31 characters were optimized unequivocally and three of them changed uniquely along the branches of the simplified molecular tree (Fig. 3). Character 9 switched from a short and thick to a long and thin style along the branch leading to *Thamnosma*/*Boenninghausenia*, character 11 switched from introrse to slightly-introrse anther opening along the same branch, and character 34 switched from the absence to the presence of acridones of type H1 along the branch subtending *Ruta*/*Thamnosma*/*Boenninghausenia* (see also Appendices 2, 3). In contrast, a transition from obovate to ovate petals (character 5) and a switch from the absence to the presence of 2-quinolones of type G1.1 (character 23) occurred more than once, but nonetheless supported the clades *Haplophyllum*/*Cneoridium* and *Thamnosma*/*Boenninghausenia*, respectively (Fig. 3).

Discussion

***Circumscription of Ruteae and* Ruta**

In the most comprehensive classification of Rutaceae, Engler (1896, 1931) proposed the inclusion of *Dictamnus* in Ruteae. However, the inferred cpDNA phylogeny strongly supports the monophyly of a clade formed by *Ruta, Boenninghausenia, Thamnosma, Haplophyllum* and *Cneoridium*, while *Dictamnus* is embedded in a clade comprising members of Zantoxyleae, Diosmeae and Toddalioideae (Fig. 1). Molecular evidence thus apparently corroborates Da Silva's (1988) interpretation of tribal boundaries (see Table 1).

The five genera sharing a single origin in the molecular tree (i.e., *Ruta*, *Boenninghausenia*, *Thamnosma*, *Haplophyllum* and *Cneoridium*; Ruteae s.s. from now on; see Fig. 1) share a number of morphological and phytochemical traits, including: the presence of actinomorphic, creamy-white to bright-yellow flowers (Engler 1896, 1931); the highest levels of lignans of the aryltetrahydronaphthalene type in Rutaceae (Waterman, 1983; Da Silva *et al*., 1988); specific classes of coumarins and acridones (Waterman, 1975); and, uniquely in Rutaceae, the biosynthetic pathway for acridones devoid of an oxygen substituent at the C-3 position and also, in some cases, at the C-1 position (Waterman, 1983). Therefore, these genera appear to form a cohesive taxonomic group.

At least morphologically, *Dictamnus* can be viewed as an aberrant form of uncertain phylogenetic placement, for it shows some morphological features that cannot be readily linked with any other taxa of Rutaceae (Moore, 1936). Within Ruteae (Engler, 1896, 1931), *Dictamnus* differs from other genera in seed structure (Corner, 1976) and floral morphology, with large, zygomorphic, white to purple flowers, lanceolate petals, and unusual oil glands protruding from the carpel walls and the style (Moore, 1936; Gut, 1966). The secondary chemistry of the genus is also unique within Rutaceae. *Dictamnus* has limonoids, instead of coumarins, and special quinolones (Da Silva *et al*., 1988). Furthermore, early serodiagnostic studies on Rutaceae (Bärner, 1927) showed that "between *Ruta chalepensis* and *Dictamnus albus* the reaction was only slightly positive, an observation strictly in accord with floral anatomy, but disagreeing with the taxonomists' assignment of *Ruta* and *Dictamnus* near to one another." (Moore, 1936: 322)

Despite the morphological and phytochemical differences between *Dictamnus* and the other genera of Ruteae and its distant relationship with other Ruteae in the cpDNA phylogeny (Fig. 1), the S-H test suggests that this phylogenetic result is not significantly different than the inclusion of *Dictamnus* in Ruteae (Table 3, constraint i). This result contrasts with the significant rejection of the second constraint assessed by the S-H test, i.e., forcing *Haplophyllum* to be sister to *Ruta* (see below). However, it should be noted that forcing *Dictamnus* in Ruteae applies a rather relaxed constraint compared to forcing *Haplophyllum* to be sister to *Ruta*, because in the former case all the trees with all possible placements of *Dictamnus* within Ruteae are allowed, whereas in the latter case only trees where *Haplophyllum* is sister to *Ruta* are allowed. Therefore, it is difficult to compare the significance values of intrinsically different constraints. To our knowledge, the potential influence of the stringency of the constraint on the significance levels estimated by the S-H test has not yet been investigated. Hence, while the results of the S-H test do not reject the possible inclusion of *Dictamnus* in Ruteae, the optimal cpDNA tree topology, morphological

observations (Moore, 1936; Gut, 1966; Corner, 1976), and phytochemical data (Da Silva *et al*., 1988) all suggest that this genus may not have evolved from the same common ancestor as the other members of Ruteae (Engler, 1896, 1931).

In Engler's (1896, 1931) classification, *Ruta* comprised around 60 species, subdivided in subgenus *Euruta*, with five species, and subgenus *Haplophyllum*, with about 50 species (Table 1). If Engler's interpretation were correct, species ascribed to the two subgenera would be expected to form sister clades in a phylogeny. However, this is not the case in the cpDNA tree inferred in this study (Fig. 1), for the eight currently-recognized species of *Ruta* (*R. chalepensis, R. angustifolia, R. graveolens, R. corsica, R. montana, R. pinnata, R. microcarpa,* and *R. oreojasme*; Townsend, 1968; Bramwell and Bramwell, 2001) form a strongly supported monophyletic group that is sister to *Boenninghausenia*/*Thamnosma*, while the 24 species of *Haplophyllum* constitute the sister clade of *Cneoridium*. Moreover, when *Haplophyllum* was constrained to be sister to *Ruta,* the differences between the likelihood score of the constrained and unconstrained topologies, evaluated by means of the S-H test, were statistically significant (Table 3). Hence, the molecular results corroborate Townsend's (1986) interpretation of generic boundaries (Table 1).

Townsend (1986) identified morphological differences between *Ruta* and *Haplophyllum* that had been overlooked by Engler (1896, 1931; Table 2). In fact, the petal margins of *Ruta* are dentate or fimbriate, whereas those of *Haplophyllum* are more or less entire; in *Ruta* the lateral flowers are 4-merous and the terminal flowers are 5-merous, whereas in *Haplophyllum* both lateral and terminal flowers have the same merism (usually 5-merous); the seeds of *Ruta* are bluntly- to sharply-angled dorsally, whereas they are reniform and dorsally-rounded in *Haplophyllum*; and finally the pollen grains of *Ruta* have elongate costae and a finely reticulate to perforate tectum ornamentation, whereas the pollen grains of *Haplophyllum* have thick costae and a closed-striate tectum. Phytochemically, while *Haplophyllum* has a predominance of alkaloids over coumarins (45%), all its most closely related genera have either more coumarins (*Ruta*, 86%; *Thamnosma*, 88%; *Boenninghausenia*, 90%) or exclusively coumarins (*Cneoridium*), an observation that, combined with additional phytochemical evidence, led Da Silva *et al*. (1988) to recommend that *Haplophyllum* be recognized as a distinct genus from *Ruta*. Therefore, our molecular phylogenetic results, statistical tests on constrained topologies, Townsend's (1986) detailed morphological analyses, and Da Silva *et al*. (1988) phytochemical data all suggest that *Haplophyllum* should be treated as a distinct genus, rather than as a subgenus of *Ruta*.

Within *Ruta*, the seven species represented by multiple accessions were all supported as monophyletic by the cpDNA genome and the main clades are congruent with

morphological observations or distribution (see Fig. 1). For example, San Miguel (2003) stated that *R. angustifolia*, *R. chalepensis*, and *R. graveolens* are morphologically very similar and virtually impossible to differentiate in their vegetative parts, whereas *R. montana* is distinguished by its narrower leaves. The three species of *Ruta* from the Canary Islands are distinct from the remaining species of the genus by being taller (Townsend, 1968; Bramwell and Bramwell, 2001) and having larger leaves (G. Salvo, personal observation), consistent with the observation that insular species exhibit trends toward larger size (Lomolino *et al*., 2006).

Recently, the populations of *R. corsica* from Sardinia were described as a new species, *R. lamarmorae*, distinguished from the Corsican populations of *R. corsica* (i.e., *R. corsica s. str*.) by morphological, ecological and karyological differences (Bacchetta *et al*., 2006). *R. corsica s. str.* is diploid, has smaller flowers, stamens, and ovaries, occurs across a wider altitudinal range (1000-1900 m.a.s.l.) and its pollen matures in June. *R. lamarmorae* is tetraploid, has bigger flowers, stamens and ovaries, occurs in a more restricted altitudinal range (1500-1750 m.a.s.l.), and its pollen matures in May. In the molecular phylogeny (Fig. 1), the accessions from Corsica and Sardinia formed two strongly supported clades, rather than being interspersed, thus suggesting that the treatment of *R. lamarmorae* as a separate species might be warranted. The phylogenetic separation between *R. corsica s. str.* and *R. lamarmorae* reflects the comparatively high absolute number (seven; data not shown) of nucleotide substitutions between the respective accessions from Corsica and Sardinia, two islands in close proximity to each other. In contrast, the accessions of *R. chalepensis* from distant locations (Sicily, Greece, Corsica, Sardinia, mainland France; see Appendix 1) are separated at most by five nucleotide substitutions (data not shown). A more definitive assessment of the proposed specific rank of *R. lamarmorae* (Bacchetta *et al*., 2006) must await further evidence from the nuclear genome, molecular dating analyses, and inter-fertility studies.

Comparisons among different datasets

Phylogenies of *Ruta* and related genera inferred from six-taxon morphological, phytochemical, and molecular matrices were compared to identify and localize incongruence between the datasets (Fig. 2), rather than to argue for or against combining data (e.g., Cannatella *et al*., 1998). Nodes with low statistical support ambiguously represent hierarchical patterns within individual datasets, thus conflict among datasets cannot be inferred from comparisons involving weak nodes (de Queiroz, 1993; Mason-Gamer and Kellogg, 1996; Graham *et al*., 1998; Van der Niet *et al*., 2005). Following this logic, direct node-to-node comparisons of bootstrap values detected no incongruent relationships between the molecular

and morphological trees (Fig. 2a, b), while the ILD test suggested significant incongruence. Beyond the known criticisms of the ILD test (e.g., Dolphin *et al*., 2000; Yoder *et al*., 2001; Darlu and Lecointre, 2002; Quicke *et al*., 2007), in the case of the molecular and morphological trees of *Ruta* and related genera, the incongruence estimated by the ILD test might reflect sampling bias in the smaller morphological dataset (12 informative characters, as compared to the 179 informative characters of the molecular dataset; Table 5), rather than conflicting phylogenetic signals (e.g., Cannatella *et al*., 1998; Graham *et al*., 1998). The only significant inconsistencies among the three trees (Fig. 2a, b, c) involved the relationships of *Ruta* and *Haplophyllum*, for the two genera were sister to each other in the phytochemical tree (BP 97), but in the molecular tree the former was sister to *Boenninghausenia*/*Thamonosma* (BP 100) and the latter to *Cneoridium* (BP 99). What could cause the observed incongruence between the molecular and phytochemical topologies?

In the phylogeny generated from parsimony analysis of the six-taxon cpDNA matrix (Fig. 2a), the terminal branches subtending *Haplophyllum* and *Cneoridium* are the longest, with 147 and 85 steps, respectively, while the branch leading to their common ancestor is much shorter (38 steps; results not shown). Since parsimony methods are known to be particularly vulnerable to long-branch attraction (LBA), as compared to model-based methods (Felsenstein, 1978; Hendy and Penny, 1989; Lewis, 2001), it is possible that the *Haplophyllum*/*Cneoridium* clade in the molecular tree (Fig. 2a) be a product of LBA. However, the analysis of the 73-accession molecular matrix by either parsimony or Bayesian methods (Fig. 1) supported the sister relationship between *Haplophyllum* and *Cneoridium*, thus suggesting that LBA should not have biased our results for the six-taxon data set.

Homoplasy is often invoked to explain disagreements among phylogenies inferred from different types of data (e.g., Sanderson and Hufford, 1996; Wiens *et al*., 2003). In secondary chemical compounds, homoplasy has been repeatedly documented, because similar selective pressures can lead to the evolution of pathways producing similar end-products in unrelated taxa (Price, 1963). Within Rutaceae, the expression of coumarin prenylation patterns, furoquinoline and acridone oxygenation patterns, and the development of the acridone and carbazole nuclei are all known to occur in unrelated taxa (Waterman, 1990). Most of the secondary chemical compounds used to infer the phytochemical phylogeny of Fig. 2c are alkaloids. Some studies have shown that alkaloid biosynthesis in plants is both plastic and labile, for the responsible genes can be repeatedly switched on and off during development and across evolutionary times (McKey, 1980; Wink and Witte, 1983; Waterman, 1998; Wink, 2003). For example, the genes responsible for the biosynthesis of quinolizidine alkaloids are widely represented in the plant kingdom, but are only expressed in a few

unrelated families (Wink and Witte, 1983). Similar examples include the evolution of the benzylisoquinoline alkaloid biosynthesis (Liscombe *et al*., 2005), and the production of pyrrolizidine alkaloids as a defence against herbivores (Reimann *et al*., 2004). Waterman (1975) commented that the alkaloid types found in Rutaceae, with the exception of the canthinones and carbazoles, exhibit a highly random distribution and fail to support any of the taxonomic groups proposed by Engler (1896, 1931). Therefore, considering the well-known problem of convergence among secondary compounds (Hegnauer, 1966; Mothes, 1981), it seems more likely that the incongruent placement of *Haplophyllum* in the phytochemical and molecular trees (Fig. 2a, c) might be caused by convergence of similar alkaloids in *Haplophyllum* and *Ruta*, rather than long-branch attraction between *Haplophyllum* and *Cneoridium*.

The analysis of phytochemical data is hampered not only by problems of homoplasy, but also by methodological complications, for example: (i) a specific compound can be detected only when a sufficient amount of it is present in the plant (Price, 1963); and (ii) chemotaxonomic reports for Rutaceae rarely mention the names of the species for which certain compounds were sought but were not detected, a problem also encountered in an angiosperm-wide study (Nandi *et al*., 1998). Therefore, the absence of a specific compound from certain species in a phytochemical data matrix may indicate true absence, lack of detection due to technical limitations, or missing information. Hence, despite the considerable diversity of secondary chemical compounds in Rutaceae (e.g., Price, 1963; Waterman 1975, 1983, 1990; Kong *et al*., 1986; Da Silva *et al*., 1988; Samuel *et al*., 2001; Zakaria, 2001), their systematic and phylogenetic value appears to be fundamentally flawed by homoplasy and methodological issues. Considering the above-mentioned problems, Waterman (1998: 547), the foremost expert on the secondary compounds of Rutaceae (Fish and Waterman, 1973; Waterman 1975, 1983, 1990; Gray and Waterman, 1978; Waterman and Grundon, 1983), asserted that "chemical systematics remains as much an art as a science, and the most appropriate use of chemical data appears to be to test phylogenies that have arisen from the interpretation of more complete non chemical datasets."

Character mapping analysis and genus diagnosis

The choice of the characters used to build the phylogenies that are in turn utilized to analyze the evolution of selected character sets remains a controversial issue, because, at its core, it influences assessments of homology and homoplasy in different data sets (Brooks, 1996). Considering the dearth, overlapping nature, and conflicting taxonomic value of the morphological characters traditionally used for Ruteae classifications (Townsend, 1986) and the well-known problem of convergence in the phytochemical data of Rutaceae (Waterman,

1990), it seems reasonable to use the strongly supported topology of the molecular tree (Fig. 1, 2a) for the mapping of non-molecular characters (Fig. 3, 4).

The optimization of character-state transitions on the molecular topology underscored the difficulty of finding non-molecular synapomorphies that are consistent with the clades of the molecular tree. Out of 47 mapped characters, only five provided character-state transitions that supported the molecular clades (characters 5, 9, 11, 23, 34; Fig. 3). Moreover, the morphological characters used by Engler (1896; 1931) to differentiate among the genera of Ruteae (i.e., characters 1, 2, 3, 4, 7, 8, 12, 14; Appendix 2) were either equivocally reconstructed or inconsistent with the molecular clades. Engler (1896, 1931) placed high taxonomic importance on flower merism (Table 2). However, the numbers of sepals, petals, stamens, and carpels are polymorphic in *Ruta* and the latter also in *Haplophyllum* (characters 1, 2, 3, 4, see Appendix 2; Saunders, 1934; Moore, 1936). Additionally, all four characters were equivocally reconstructed (Fig. 4). Consequently, the change between five and four elements in the perianth whorls and between ten and eight stamens in the androecium could have occurred either by reduction or by expansion (Fig. 4), corroborating the suggestion that these characters may be evolutionarily labile (Endress, 1990, 1999), and thus of limited taxonomic value. Conversely, in the gynoecium, the transition to two carpels in *Thamnosma* and one in *Cneoridium* always occurred by reduction, regardless of the ancestral state for the clade formed by *Ruta*, *Boenninghausenia*, *Thamnosma*, *Haplophyllum*, and *Cneoridium* (Ruteae s.s.; Fig. 4). Similar reductionist trends have been previously reported in other angiosperm families and explained in terms of paedomorphic development caused by the elimination of the last initiated organ (Tucker *et al*., 1993; Hufford, 1996). An alternative interpretation, also consistent with the optimizations of floral merism on the molecular topology (Fig. 4), assumes that the ancestor of Ruteae s.s. was polymorphic for floral merism, thus implying that *Ruta* retained the ancestral polymorphisms for all floral whorls and *Haplophyllum* for the gynoecium only. Fixation of the number of organs in the perianth and androecium occurred in the lineages leading to *Boenninghausenia*/*Thamnosma* and *Cneoridium/Haplophyllum*, respectively, while carpel number was reduced in the ancestor of *Boenninghausenia* and *Thamnosma* and, independently, in *Cneoridium* (Fig. 4). Repeated processes of selection leading to fixation of character states from a polymorphic ancestor have indeed been proposed as a likely evolutionary explanation for multiple transitions between character states (Brooks, 1996). Thus the results of character mapping indicate that the way forward to understand patterns of homoplasy in the floral morphology of Ruteae *s.s*. may lie in detailed comparisons between the ontogenetic trajectories of floral whorls.

Within the context of the phylogenetic relationships supported by molecular data (Figs. 1 and 2a), *Ruta*, the type genus for the family, can be diagnosed by using a combination of plesiomorphic, homoplasious, and autapomorphic morphological character states, including: obovate petals and short, thick style (both plesiomorphic states, for they are also present in the outgroup); cushion-shaped nectary disk, elongate anthers, and uncurved embryo (all homoplasious states present also in *Haplophyllum*); introrse anthers' opening (a homoplasious state present also in *Haplophyllum* and *Cneoridium*); simple stigma shape and seeds angled dorsally (two homoplasious states present also in *Boenninghausenia* and *Thamnosma*, respectively); and dentate or fimbriate petal margins (an autapomorphy for *Ruta*; see Appendix 2).

Conclusion

The finding that traditionally important taxonomic characters do not provide support for the clades identified in molecular phylogenies has become a frequent occurrence with the widespread development of molecular systematics (e.g., Lavin *et al*., 2001; Moylan *et al*., 2004; Van der Niet *et al*., 2005; Norup *et al*., 2006). When mapped onto a molecular phylogeny, characters originally used to build classifications have been found to be plesiomorphic (Lavin *et al*., 2001), homoplasious (Moylan *et al*., 2004; Norup *et al*., 2006), or simply uninformative for diagnosing clades (Van der Niet *et al*., 2005). It has been remarked that the frequency of homoplasy in traditional taxonomic characters may reflect the fact that they are usually optimized on molecular topologies, thus establishing an *a priori* bias for homoplasy in non-molecular data sets (Grant, 2003). In other words, homoplasy may in part derive from inappropriate comparisons between classes of characters at different hierarchical levels of organization (Doyle, 1996; Minelli, 1998). This consideration may be especially relevant to the homoplasy observed in the mapping of phytochemical characters onto the Ruteae molecular tree, for different biogenetic pathways can produce the same compound (Hegnauer, 1966; Waterman, 1990). Consequently, the biogenetic pathways leading to these compounds, and not the compounds *per se*, should be examined and scored as character states. Rather than being viewed in a negative sense, the identification of homoplasy in non-molecular characters should be used as a starting point to study its biological and methodological causes, focusing especially on the developmental pathways underlying phenotypic traits and the conflicting assessments of homology often performed when comparing characters at different levels of organization.

Acknowledgments

The authors would like to thank Víctor Suárez Santiago (*Haplophyllum* from Spain), Oyumaa Jamsran (*H. dauricum* from Mongolia), Louis Zeltner (*H.* from Iran and help in the field), Laszlo Csiba, Vincent Savolainen and Mark Chase (Rutaceae from the Kew Gardens DNA Bank), Jesús Muñoz, Javier Fuertes Aguilar, Pablo Vargas, and Gonzalo Nieto Feliner (help and material from the Real Jardín Botánico of Madrid), Dennis Hansen (useful contacts), Angel B. Fernández (*Ruta microcarpa* leaves), Alfredo Valido Amador (*R. oreojasme* and *R. pinnata* leaves), Peter Endress (help with the German literature and useful conversations throughout the research), Rolf Rutishauser (help on Rutaceae morphology), Mary Endress (material from the Missouri Botanical Gardens and comments on an early draft), Andre Simoes (material from Meise and help in the lab), Daniel Heinzmann (help in the lab), Guilhem Mansion (help in the field and during the research), Timo van der Niet, and Chloe Galley (help with the analyses), Brigitte Marazzi, Alessia Guggisberg, and Barbara Keller (help throughout the research and comments on an early draft), Elena Benetti and Barbara Seitz (precious help in the library). This study was funded by the Swiss National Science Foundation grant no. 3100A0-105495 to E.C. and the Institute of Systematic Botany of the University of Zurich. Additional financial support from the Georges-und-Antoine-Claraz-Schenkung, the Kommission für Reisestipendien der Schweizerischen Akademie der Naturwissenschaften (SANW), and the Equal Opportunity Office of the University of Zurich are gratefully acknowledged.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Contrology 19, 716-723.

Bacchetta, G., Brullo, S., Giusso del Galdo, G., 2006. *Ruta lamarmorae* (Rutaceae), a new species from Sardinia. Edinburgh Journal of Botany 63, 153-160.

Baker, R.H., Yu, X.B., DeSalle, R., 1998. Assessing the relative contribution of molecular and morphological characters in simultaneous analysis trees. Molecular Phylogenetics and Evolution 9, 427-436.

Baum, D.A., Small, R.L., Wendel, J.F., 1998. Biogeography and floral evolution of baobabs (*Adansonia*, Bombacaceae) as inferred from multiple data sets. Syst. Biol. 47, 181-207.

Bärner, J., 1927. Serodiagnostische Verwandschaftforschungen innerhalb der Geraniales, Sapindales, Rhamnales und Malvales. Bibl. Bot., Stuttgart, 94, 1-39.

Bramwell, D., Bramwell, Z.I., 2001. Wild flowers of the Canary Islands. Editorial Rueda, Madrid, Spain.

Brooks, D.R., 1996. Explanations of homoplasy at different levels of biological organization. In: Sanderson, M.J., Hufford, L. (Eds.), Homoplasy: the recurrence of similarity in evolution. Academic Press, San Diego, pp. 3-36.

Bull, J.J., Huelsenbeck, J.P., Cunningham, C.W., Swofford, D.L., Waddell, P.J., 1993. Partitioning and combining data in phylogenetic analysis. Syst. Biol. 42, 384-397.

Cannatella, D.C., Hillis, D.M., Chippindale, P.T., Weigt, L., Rand, A.S., Ryan, M.J., 1998. Phylogeny of frogs of the *Physalaemus pustulosus* species group, with an examination of data incongruence. Syst. Biol. 47, 311-335.

Chase, M.W., Morton, C.M., Kallunki, J.A., 1999. Phylogenetic relationships of Rutaceae: a cladistic analysis of the subfamilies using evidence from *rbc*L and *atp*B sequence variation. Am. J. Bot. 86, 1191-1199.

Corner, E.J.H., 1976. The seeds of Dicotyledons. Cambridge University Press, Cambridge.

Cunningham, C.W., 1997. Is congruence between data partitions a reliable predictor of phylogenetic accuracy? Empirically testing an iterative procedure for choosing among phylogenetic methods. Syst. Biol. 46, 464-478.

Da Silva, M.F., Gottlieb, O.R., Ehrendorfer, F., 1988. Chemosystematics of the Rutaceae: suggestions for a more natural taxonomy and evolutionary interpretation of the family. Plant Syst. Evol. 161, 97-134.

Darlu, P., Lecointre, G., 2002. When does the Incongruence Length Difference test fail? Mol. Biol. Evol. 19, 432-437.

de Queiroz, A., 1993. For consensus (sometimes). Syst. Biol. 42, 368-372.

de Queiroz, A., Donoghue, M.J., Kim, J., 1995. Separate versus combined analysis of phylogenetic evidence. Annu. Rev. Ecol. Syst. 26, 657-681.

Dolphin, K., Belshaw, R., Orme, C.D.L., Quicke, D.L.J., 2000. Noise and incongruence: interpreting results of the Incongruence Length Difference test. Molecular Phylogenetics and Evolution 17, 401-406.

Donoghue, M.J., Sanderson, M.J., 1992. The suitability of molecular and morphological evidence in reconstructing plant phylogeny. In: Soltis, P., Soltis, D., Doyle, J. (Eds.), Molecular systematics of plants. Chapman and Hall, New York, London, pp. 340-368.

Doyle, J.J., 1996. Homoplasy connections and disconnections: genes and species, molecules and morphology. In: Sanderson, M.J., Hufford, L. (Eds.), Homoplasy: the recurrence of similarity in evolution. Academic Press, San Diego, pp. 37-66.

Doyle, J.J., Doyle, J.L., Rauscher, J.T., Brown, A.H.D., 2004. Diploid and polyploid reticulate evolution throughout the history of the perennial soybeans (*Glycine* subgenus *Glycine*). New Phytologist 161, 121-132.

Endress, P.K., 1990. Patterns of floral construction in ontogeny and phylogeny. Biological Journal of the Linnaean Society 39, 153-175.

Endress, P.K., 1999. Symmetry in flowers: diversity and evolution. Int. J. Plant Sci. 160, S3-S23.

Engler, A., 1896. Rutaceae. In: Engler, A., Prantl, K. (Eds.), Nat. Pflanzenfam III, Vol. 4. Engelmann, Leibzig, pp. 95-201.

Engler, A., 1931. Rutaceae. In: Engler, A., Prantl, K. (Eds.), Die natürlichen Pflanzenfamilien, 2. Aufl. 19a. Engelmann, Leibzig, pp. 187-359.

Farris, J.S., Kallersjo, M., Kluge, A.G., Bult, C., 1994. Testing significance of incongruence. Cladistics 10, 315-319.

Felsenstein, J., 1978. Cases in which parsimony or compatibility methods will be positively misleading. Systematic Zoology 27, 401-410.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the Bootstrap. Evolution 39, 783-791.

Fish, F., Waterman, P.G., 1973. The chemosystematics of the *Zanthoxylum*/*Fagara* complex. Taxon 22, 177-203.

Fitch, W.M., 1971. Towards defining the course of evolution: minimum change for a specific tree topology. Systematic Zoology 20, 406-416.

Gadek, P.A., Fernando, E.S., Quinn, C.J., Hoot, S.B., Terrazas, T., Sheahan, M.C., Chase, M.W., 1996. Sapindales: molecular delimitation and infraordinal groups. Am. J. Bot. 83, 802-811.

Gatesy, J., Baker, R.H., 2005. Hidden likelihood support in genomic data: can forty-five wrongs make a right? Syst. Biol. 54, 483-492.

Goldman, N., Anderson, J.P., Rodrigo, A.G., 2000. Likelihood-based tests of topologies in phylogenetics. Syst. Biol. 49, 652-670.

Graham, S.W., Kohn, J.R., Morton, B.R., Eckenwalder, J.E., Barrett, S., 1998. Phylogenetic congruence and discordance among one morphological and three molecular data sets from Pontederiaceae. Syst. Biol. 47, 545-567.

Grant, V., 2003. Incongruence between cladistic and taxonomic systems. Am. J. Bot. 90, 1263-1270.

Gray, A.I., Waterman, P.G., 1978. Coumarins of the Rutaceae. Phytochemistry 17, 845-864.

Guggisberg, A., Mansion, G., Kelso, S., Conti, E., 2006. Evolution of biogeographic patterns, ploidy levels, and breeding systems in a diploid-polyploid species complex of *Primula*. New Phytologist 171, 617-632.

Gut, B.J., 1966. Beiträge zur morphologie des gynoeceums und der blütenachse einiger Rutaceen. Bot. Jahrb. 85, 10-247.

Hart, M.W., Johnson, S.L., Addison, J.A., Byrne, M., 2004. Strong character incongruence and character choice in phylogeny of sea stars of the Asterinidae. Invertebr. Biol. 123, 343-356.

Hegnauer, R., 1966. Comparative phytochemistry of alkaloids. In: Swain, T. (Ed.), Comparative phytochemistry. Academic Press, London and New York, pp. 211-230.

Hendy, M.D., Penny, D., 1989. A framework for the quantitative study of evolutionary trees. Systematic Zoology 38, 297-309.

Hillis, D.M., 1987. Molecular versus morphological approaches to systematics. Annu. Rev. Ecol. Syst. 18, 23-42.

Hillis, D.M., 1995. Approaches for assessing phylogenetic accuracy. Syst. Biol. 44, 3-16.

Hufford, L., 1996. Ontogenetic evolution, clade diversification, and homoplasy. In: Sanderson, M.J., Hufford, L. (Eds.), Homoplasy: the recurrence of similarity in evolution. Academic Press, San Diego, pp. 271-301.

Hughes, C.E., Lewis, G.P., Yomona, A.D., Reynel, C., 2004. *Maraniona*. A new dalbergioid legume genus (Leguminosae, Papilionoideae) from Peru. Syst. Bot. 29, 366-374.

Hulsenbeck, J.P., Ronquist, F., 2001. Mr Bayes: Bayesian inference of phylogeny. Bioinformatics 17, 754-755.

Jussieu, A., 1825. Rutaceae: *Haplophyllum*. Mém. Mus. Hist. Nat. 12, 464.

Kelchner, S.A., 2000. The evolution of non-coding chloroplast DNA and its application in plant systematics. Ann. Missouri Bot. Gard. 87, 482-498.

Kennedy, M., Holland, B.R., Gray, R.D., Spencer, H.G., 2005. Untangling long branches: identifying conflicting phylogenetic signals using Spectral Analysis, Neighbor-Net, and Consensus Networks. Syst. Biol. 54, 620-633.

Kim, S.-C., Crawford, D.J., Jansen, R.K., 1996. Phylogenetic relationships among the genera of the subtribe Sonchinae (Asteraceae): evidence from ITS sequences. Syst. Bot. 21, 417-432.

Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol. 29, 170-179.

Kluge, A.G., 1998. Total evidence or taxonomic congruence: cladistics or consensus classification. Cladistics 14, 151-158.

Kong, Y.-C., Cheng, K.-F., Ng, K.-M., But, P.P.-H., Li, Q., Yu, S.-X., Chang, H.-T., Cambie, R.C., Kinoshita, T., Kan, W.-S., Waterman, P.G., 1986. A chemotaxonomic division of *Murraya* based on the distribution of the alkaloids yuehchukene and girinimbine. Biochem. Syst. Ecol. 14, 491-497.

Kron, K.A., Judd, W.S., Crayn, D.M., 1999. Phylogenetic analyses of Andromedeae (Ericaceae subfam. Vaccinioideae). Am. J. Bot. 86, 1290-1300.

Lavin, M., Pennington, R.T., Klitgaard, B.B., Sprent, J.I., de Lima, H.C., Gasson, P.E., 2001. The dalbergioid legumes (Fabaceae): delimitation of a pantropical monophyletic clade. Am. J. Bot. 88, 503-533.

Lee, M.S.Y., 2001. Uninformative characters and apparent conflict between molecules and morphology. Mol. Biol. Evol. 18, 676-680.

Lee, M.S.Y., Hugall, A.F., 2003. Partitioned Likelihood Support and the evaluation of data set conflict. Syst. Biol. 52, 15-22.

Lewis, P.O., 2001. A Likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50, 913-925.

Linnaeus, C., 1735. Systema Naturae, sive regna tria naturae, systematice proposita per classes, ordines, genera, et species. Apud Theodorum Haak, Leiden.

Linnaeus, C., 1753. Species Plantarum, exhibentes plantas rite cognitas, ad genera relatas, cum differentiis specificis, nominibus trivialibus, synonymis selectis, locis natalibus, secundum systema sexuale digestas. Holmiae, Impensis Laurentii Salvii, Stockholm.

Liscombe, D.K., MacLeod, B.P., Loukanina, N., Nandi, O.I., Facchini, P.J., 2005. Evidence for the monophyletic evolution of benzylisoquinoline alkaloid biosynthesis in angiosperms. Phytochemistry 66, 2500-2520.

Lomolino, M.V., Riddle, B.R., Brown, J.H., 2006. Biogeography, 3rd ed. Sinauer Associates, Sunderland, USA.

Mabberley, D.J., 2002. The plant-book: a portable dictionary of the vascular plants, 2nd ed. Cambridge University Press, Cambridge.

Maddison, P.G., Maddison, D.R., 2000. MacClade4: analysis of phylogeny and character evolution. Sinauer Associates, Sunderland, USA.

Marazzi, B., Endress, P.K., De Queiroz, L.P., Conti, E., 2006. Phylogenetic relationships within *Senna* (Leguminosae, Cassiinae) based on three chloroplast DNA regions: patterns in the evolution of floral symmetry and extrafloral nectaries. Am. J. Bot. 93, 288-303.

Mason-Gamer, R.J., Kellogg, E.A., 1996. Testing for phylogenetic conflict among molecular data sets in the Triticeae (Gramineae). Syst. Biol. 45, 524-545.

McKey, D., 1980. Origins of novel alkaloid types: a mechanism for rapid phenotypic evolution of plant secondary compounds. Am. Nat. 115, 754-759.

Mester, I., Vicol, E.C., 1971. Contribution to the chemotaxonomy of the *Haplophyllum* genus. Rev. Roum. Biol. 16, 221-233.

Minelli, A., 1998. Molecules, developmental modules, and phenotypes: a combinatorial approach to homology. Molecular Phylogenetics and Evolution 9, 340-347.

Miyamoto, M., Fitch, W.M., 1995. Testing species phylogenies and phylogenetic methods with congruence. Syst. Biol. 44, 64-76.

Moore, J.A., 1936. Floral anatomy and phylogeny in the Rutaceae. New Phytol. 35, 318-322.

Morton, C.M., Grant, M., Blackmore, S., 2003. Phylogenetic relationships of the Aurantioideae inferred from chloroplast DNA sequence data. Am. J. Bot. 90, 1463-1469.

Mothes, K., 1981. The problem of chemical convergence in secondary metabolism. Sci Scientists, 323-326.

Moylan, E.C., Bennett, J.R., Carine, M.A., Olmstead, R.G., Scotland, R.W., 2004. Phylogenetic relationships among *Strobilanthes* s.l. (Acanthaceae): evidence from ITS nrDNA, trnL-F cpDNA, and morphology. Am. J. Bot. 91, 724-735.

Mueller, R.L., Robert Macey, J., Jaekel, M., Wake, D.B., Boore, J.L., 2004. Morphological homoplasy, life history evolution, and historical biogeography of plethodontid salamanders inferred from complete mitochondrial genomes. PNAS 101, 13820-13825.

Muellner, A.N., Vassiliades, D.D., Renner, S.S., In press. Placing Biebersteiniaceae, a herbaceous clade of Sapindales, in a temporal and geographic context. Plant Syst. Evol.

Nandi, O.I., Chase, M.W., Endress, P.K., 1998. A combined cladistic analysis of angiosperms using rbcL and non-molecular data sets. Ann. Missouri Bot. Gard. 85, 137-214.

Navarro, F.B., Suarez-Santiago, V.N., Blanca, G., 2004. A new species of *Haplophyllum* A. Juss (Rutaceae) from the Iberian peninsula: evidence from morphological, karyological and molecular analyses. Ann. Bot. 94, 571-582.

Ng, K.M., But, P.P.-H., Gray, A.I., Hartley, T.G., Kong, Y.-C., Waterman, P.G., 1987. The biochemical systematics of *Tetradium*, *Euodia*, and *Melicope* and their significance in the Rutaceae. Biochem. Syst. Ecol. 15, 587-593.

Norup, M.V., Dransfield, J., Chase, M.W., Barfod, A.S., Fernando, E.S., Baker, W.J., 2006. Homoplasious character combinations and generic delimitation: a case study from the Indo-Pacific arecoid palms (Arecaceae: Areceae). Am. J. Bot. 93, 1065-1080.

Novacek, M.J., 1994. Morphological and molecular inroads to phylogeny. In: Grande, L., Rieppel, O. (Eds.), Interpreting the hierarchy of nature: from systematic patterns to evolutionary process theories. Academic Press Inc., San Diego, pp. 85-131.

Patterson, C., 1988. Molecules and morphology in evolution: conflict or compromise? Cambridge University Press, Cambridge.

Porter, J.M., Johnson, L.A., 2000. A phylogenetic classification of Polemoniaceae. Aliso 19, 55-91.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817-818.

Price, J.R., 1963. The distribution of alkaloids in the Rutaceae. In: Swain, T. (Ed.), Chemical Plant Taxonomy. Academic Press, London and New York, pp. 429-452.

Quicke, D.L.J., Jones, O.R., Epstein, D.R., 2007. Correcting the problem of false incongruence due to noise imbalance in the Incongruence Length Difference (ILD) test. Syst. Biol. 56, 496-503.

Reimann, A., Nurhayati, N., Backenköhler, A., Oberand, D., 2004. Repeated evolution of the pyrrolizidine alkaloid-mediated defense system in separate angiosperm lineages. The Pl. Cell 16, 2772-2784.

Rieseberg, L.H., Whitton, J., Linder, C.R., 1996. Molecular marker incongruence in plant hybrid zones and phylogenetic trees. Acta Bot. Neerl. 45, 243-262.

Roalson, E.H., Boggan, J.K., Skog, L.E., Zimmer, E.A., 2005. Untangling Gloxinieae (Gesneriaceae). I. Phylogenetic patterns and generic boundaries inferred from nuclear, chloroplast, and morphological cladistic datasets. Taxon 54, 389-410.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572-1574.

Rønsted, N., Salvo, G., Savolainen, V., 2007. Biogeographical and phylogenetic origins of African fig species (*Ficus* section *Galoglychia*). Molecular Phylogenetics and Evolution 43, 190-201.

Rutschmann, F., Eriksson, T., Abu Salim, K., Conti, E., 2007. Assessing calibration uncertainty in molecular dating: the assignment of fossils to alternative calibration points. Syst. Biol. 56, 591-608.

Samuel, R., Ehrendorfer, F., Chase, M.W., Greger, H., 2001. Phylogenetic analyses of the Aurantioideae (Rutaceae) based on non-coding plastid DNA sequences and phytochemical features. Plant Biol. 3, 77-87.

San Miguel, E., 2003. Rue (*Ruta* L., Rutaceae) in traditional Spain: frequency and distribution of its medicinal and symbolic applications. Econ. Bot. 57, 231-244.

Sanderson, M.J., Hufford, L., 1996. Homoplasy: the recurrence of similarity in evolution. Academic Press, San Diego, USA.

Sanderson, M.J., Wojciechowski, M.F., Hu, J.-M., Sher Khan, T., Brady, S.G., 2000. Error, bias, and long-branch attraction in data for two chloroplast photosystem genes in seed plants. Mol. Biol. Evol. 17, 782-797.

Sang, T., Crawford, D.J., Stuessy, T.F., 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). Am. J. Bot. 84, 1120-1136.

Saunders, E.R., 1934. On carpel polymorphism. IV. Ann. Bot. 48, 643-692.

Scholz, H., 1964. Rutales. In: Melchior, H., Wedermann, E. (Eds.), A. Engler's Syllabus der Pflanzenfamilien, vol. 2, 12th ed. Bornträger, Berlin, pp. 262-277.

Scott, K.D., McIntyre, C.L., Playford, J., 2000. Molecular analyses suggest a need for a significant rearrangement of Rutaceae subfamilies and a minor reassessment of species relationships within *Flindersia*. Plant Syst. Evol. 223, 15-27.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114-1116.

Simões, A., Endress, M., van der Niet, T., Conti, E., Kinoshita, L.S., 2004. Tribal and intergeneric relationships of Mesechiteae (Apocynoideae, Apocynaceae): evidence from three noncoding plastid DNA regions and morphology. Am. J. Bot. 91, 1409-1418.

Simões, A.O., Endress, M., van der Niet, T., Kinoshita, L.S., Conti, E., 2006. Is Mandevilla (Apocynaceae, Mesechiteae) monophyletic? Evidence from five plastid DNA loci and morphology. Ann. Missouri Bot. Gard. 93, 565-591.

Song, W.-H., Li, X.-D., Li, X.-W., Huang, H.-W., Li, J.-Q., 2004. Genetic diversity and conservation strategy of *Psilopeganum sinense*, a rare species in the Three-Gorges Reservoir area. Biodiversity Science 12, 227-236.

Stace, C.A., 2005. Plant taxonomy and biosystematics - does DNA provide all the answers? Taxon 54, 999-1007.

Stevens, P.F., 2001 onwards. Angiosperm Phylogeny Website. Version 8, June 2007. http://www.mobot.org/MOBOT/research/APweb/

Swofford, D.L., 1991. When are phylogeny estimates from molecular and morphological data incongruent? In: Miyamoto, M.M., Cracraft, J. (Eds.), Phylogenetic analysis of DNA sequences. Oxford Univ. Press, New York, pp. 295-333.

Swofford, D.L., 2001. PAUP*4.0b10: Phylogenetic Analysis Using Parsimony (*and other methods). Sinauer, Sunderland, USA.

Sytsma, K.J., 1990. DNA and Morphology: inference of plant phylogeny. Trends in Ecology and Evolution 5, 104-110.

Taberlet, P., Gielly, L., Pautou, G., Bouvet, J., 1991. Universal primers for amplification of 3 non-coding regions of chloroplast DNA. Plant Mol. Biol. 17, 1105-1109.

Tang, F., Ye, Q., Yao, X., Huang, H., 2007. Isolation and characterization of microsatellite loci in *Psilopeganum sinense* Hemsl. (Rutaceae), an endangered herb endemic to Yangtze River valley. Molecular Ecology Notes, published online, doi: 10.1111/j.1471-8286.2007.01933.x.

Tilak, V.D., Nene, P.M., 1978. Floral anatomy of the Rutaceae. Ind. J. Bot. 1, 83-90.

Townsend, C.C., 1968. Rutaceae. In: Tutin T.G. (Ed.), Flora Europea, vol.2, Rosaceae to Umbelliferae. Cambridge University Press, Cambridge, pp. 227-230.

Townsend, C.C., 1986. Taxonomic revision of the genus *Haplophyllum* (Rutaceae). Hooker's Icones Plantarum, Vol. XL, parts I, II, and III. Bentham-Moxon Trustees, Kent, UK.

Tucker, S.C., Douglas, A.W., Liang, H.-X., 1993. Utility of ontogenetic and conventional characters in determining phylogenetic relationships of Saururaceae and Piperaceae (Piperales). Syst. Bot. 18, 614-641.

Van der Niet, T., Linder, P.H., Bytebier, B., Bellstedt, D.U., 2005. Molecular markers reject monophyly of the subgenera of *Satyrium* (Orchidaceae). Syst. Bot. 30, 263-274.

Wahlberg, N., Nylin, S., 2003. Morphology versus molecules: resolution of the positions of *Nymphalis*, *Polygonia*, and related genera (Lepidoptera: Nymphalidae). Cladistics 19, 213-223.

Waterman, P.G., 1975. Alkaloids of the Rutaceae: their distribution and systematic significance. Biochem. Syst. Ecol. 3, 149-180.

Waterman, P.G., 1983. Phylogenetic implications of the distribution of secondary metabolites within the Rutales. In: Waterman, P.G., Grundon, M.F. (Eds.), Chemistry and chemical taxonomy of the Rutales. Academic Press, London, pp. 377-400.

Waterman, P.G., 1990. Chemosystematics of the Rutaceae: comments on the interpretation of Da Silva *et al*. Plant Syst. Evol. 173, 39-48.

Waterman, P.G., 1998. Alkaloid chemosystematics. In: Cordell, G.A. (Ed.), The alkaloids: chemistry and biology. Academic Press, USA, pp. 537-565.

Waterman, P.G., Grundon, M.F., 1983. Chemistry and chemical taxonomy of the Rutales. Academic Press, London.

Wiens, J.J., Chippindale, P.T., Hillis, D.M., 2003. When are phylogenetic analyses misled by convergence? A case study in Texas cave salamanders. Syst. Biol. 52, 501-514.

Wiens, J.J., Fetzner, J.W., Parkinson, C.L., Reeder, T.W., 2005. Hylid frog phylogeny and sampling strategies for speciose clades. Syst. Biol. 54, 719-748.

Wiens, J.J., Hollingsworth, B.D., 2000. War of the iguanas: conflicting molecular and morphological phylogenies and Long-Branch Attraction in iguanid lizards. Syst. Biol. 49, 143-159.

Wilcox, T.P., García de León, F.J., Hendrickson, D.A., Hillis, D.M., 2004. Convergence among cave catfishes: long-branch attraction and a Bayesian relative rates test. Molecular Phylogenetics and Evolution 31, 1101-1113.

Wink, M., 2003. Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. Phytochemistry 64, 3-19.

Wink, M., Witte, L., 1983. Evidence for a wide spread occurrence of the genes of quinolizidine alkaloid biosynthesis. Induction of alkaloid accumulation in cell suspension cultures of alkaloid-"free" species. FEBS Letters 159, 196-200.

Won, H., Renner, S.S., 2003. Horizontal gene transfer from flowering plants to *Gnetum*. PNAS 100, 10824-10829.

Wortley, A.H., Scotland, R.W., 2006. Determining the potential utility of datasets for phylogeny reconstruction. Taxon 55, 431-442.

Yoder, A.D., Irwin, J.A., Payseur, B.A., 2001. Failure of the ILD to determine data combinability for slow loris phylogeny. Syst. Biol. 50, 408-424.

Young, N.D., Healy, J., 2003. Gapcoder automates the use of indel characters in phylogenetic analysis. BMC Bioinformatics 4, 6.

Zakaria, M.B., 2001. The phytochemistry of Rutaceae species with special reference to *Melicope*. Malayan Nature Journal 55, 241-250.

Figure legends

Figure 1. Ninety five % majority-rule consensus tree obtained from the Bayesian analysis. Posterior probabilities (PPs) and Bootstrap percentages (BPs; above 50%) are shown above and below branches, respectively. The white bar indicates members of tribe Ruteae sensu Engler (1896, 1931). The grey bar indicates members of tribe Zanthoxyleae. Tribes Ruteae, Zanthoxyleae, and Diosmeae belong to subfamily Rutoideae; tribe Toddalieae belongs to subfamily Toddalioideae. Taxa with an * were included in the outgroup.

Figure 2. Trees obtained from parsimony analyses of molecular (a), morphological (b), phytochemical (c), and global (d) data sets. *Choisya* served as the rooting taxon. Bootstrap percentages (BPs) generated by the same data set used to infer the tree are reported above the branches; BPs generated by the rival data sets are reported below the branches: (a) BP morphology/BP phytochemistry, (b) BP molecules/BP phytochemistry, (c) BP molecules/BP morphology.

Figure 3. The five derived non-molecular character states consistent with the clades supported by molecular data: 5, ovate petals (taken from Townsend, 1986); 9, thin and long style (taken from Engler, 1931); 11, slightly introrse anther opening (shown in cross-sectional view with the lower part facing the gynoecium and the arrows indicating the direction of pollen release; personal observation and Prof. P. Endress pers. comm.); 23, 2-quinolones of type G1.1 (taken from Da Silva *et al*., 1988); 34, acridones of type H1 (taken from Da Silva *et al*., 1988). Black and white bars indicate non-homoplasious and homoplasious transitions, respectively. *R = Ruta*, *B = Boennighausenia*, *T = Thamnosma*, *H = Haplophyllum*, *C = Cneoridium*, O = Outgroup (*Choisya*). See Appendix 2, 3 for further details.

Figure 4. Optimizations of transitions between states for the four characters associated with floral merism: number of sepals, petals, stamens, and carpels. The states of each character, representing the number of units within each whorl, are symbolized by different patterns in the boxes above of the branches; coexistence of different states for the same character within the same taxon is symbolized by a circle. Branches with horizontal lines indicate equivocal reconstructions. Name of taxa as in Fig. 3. See Appendix 2 for further details.

Figure 1

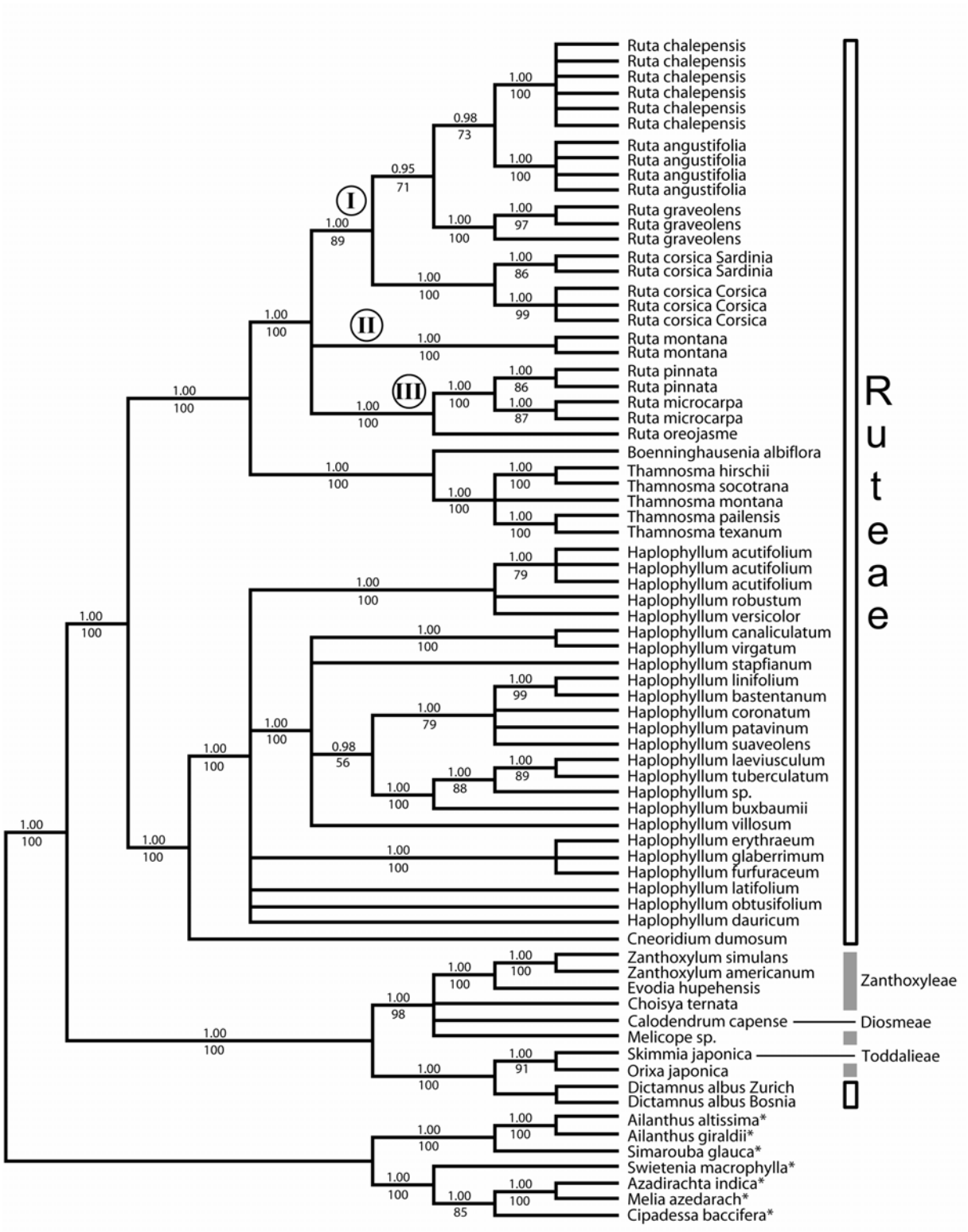
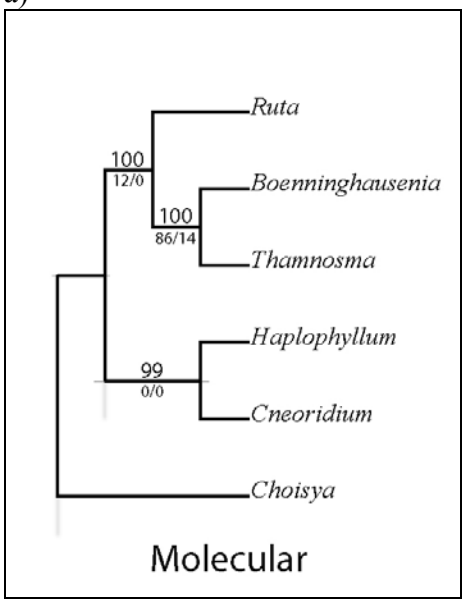
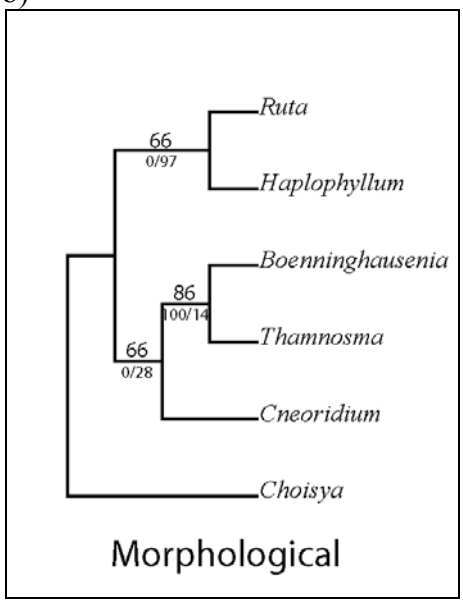

Ruta chalepensis
Ruta chalepensis
Ruta chalepensis
Ruta chalepensis
Ruta chalepensis
Ruta chalepensis
Ruta angustifolia
Ruta angustifolia
Ruta angustifolia
Ruta angustifolia
Ruta graveolens
Ruta graveolens
Ruta graveolens
Ruta corsica Sardinia
Ruta corsica Sardinia
Ruta corsica Corsica
Ruta corsica Corsica
Ruta corsica Corsica
Ruta montana
Ruta montana
Ruta pinnata
Ruta pinnata
Ruta microcarpa
Ruta microcarpa
Ruta oreojasme
Boenninghausenia albiflora
Thamnosma hirschii
Thamnosma socotrana
Thamnosma montana
Thamnosma pailensis
Thamnosma texanum
Haplophyllum acutifolium
Haplophyllum acutifolium
Haplophyllum acutifolium
Haplophyllum robustum
Haplophyllum versicolor
Haplophyllum canaliculatum
Haplophyllum virgatum
Haplophyllum stapfianum
Haplophyllum linifolium
Haplophyllum bastentanum
Haplophyllum coronatum
Haplophyllum patavinum
Haplophyllum suaveolens
Haplophyllum laeviusculum
Haplophyllum tuberculatum
Haplophyllum sp.
Haplophyllum buxbaumii
Haplophyllum villosum
Haplophyllum erythraeum
Haplophyllum glaberrimum
Haplophyllum furfuraceum
Haplophyllum latifolium
Haplophyllum obtusifolium
Haplophyllum dauricum
Cneoridium dumosum
Zanthoxylum simulans
Zanthoxylum americanum
Evodia hupehensis
Choisya ternata
Calodendrum capense
Melicope sp.
Skimmia japonica
Orixa japonica
Dictamnus albus Zurich
Dictamnus albus Bosnia
Ailanthus altissima*
Ailanthus giraldii*
Simarouba glauca*
Swietenia macrophylla*
Azadirachta indica*
Melia azedarach*
Cipadessa baccifera*
I
II
III
Rutaeae
Zanthoxyleae
Diosmeae
Toddalieae

Figure 2

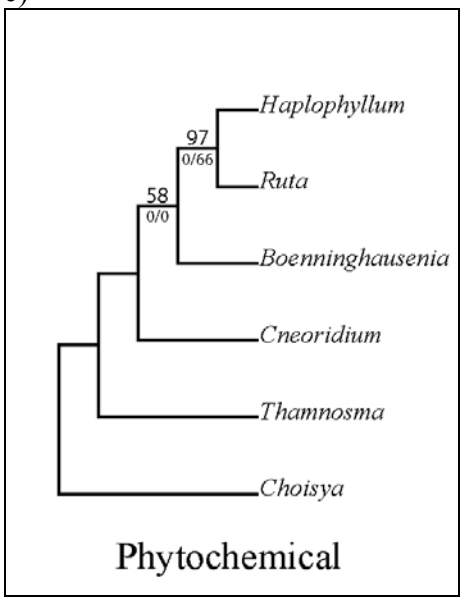
a)



b)



c)



d)

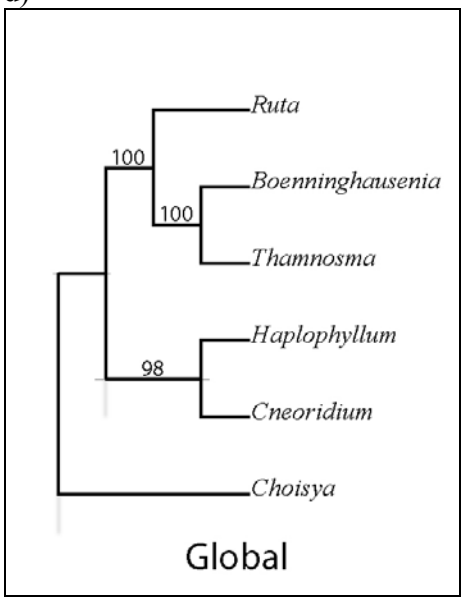


Figure 3

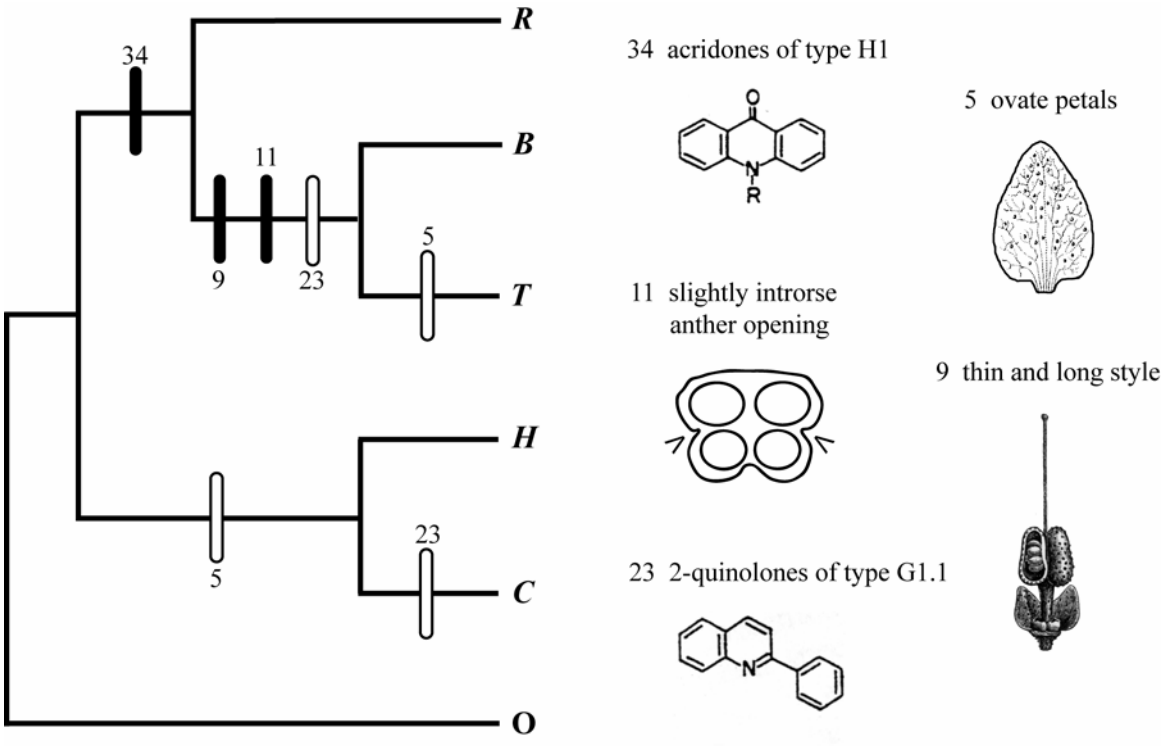

R
B
T
H
C
O
34
11
9
23
5
5
23
34 acridones of type H1
11 slightly introrse anther opening
23 2-quinolones of type G1.1
5 ovate petals
9 thin and long style

Figure 4

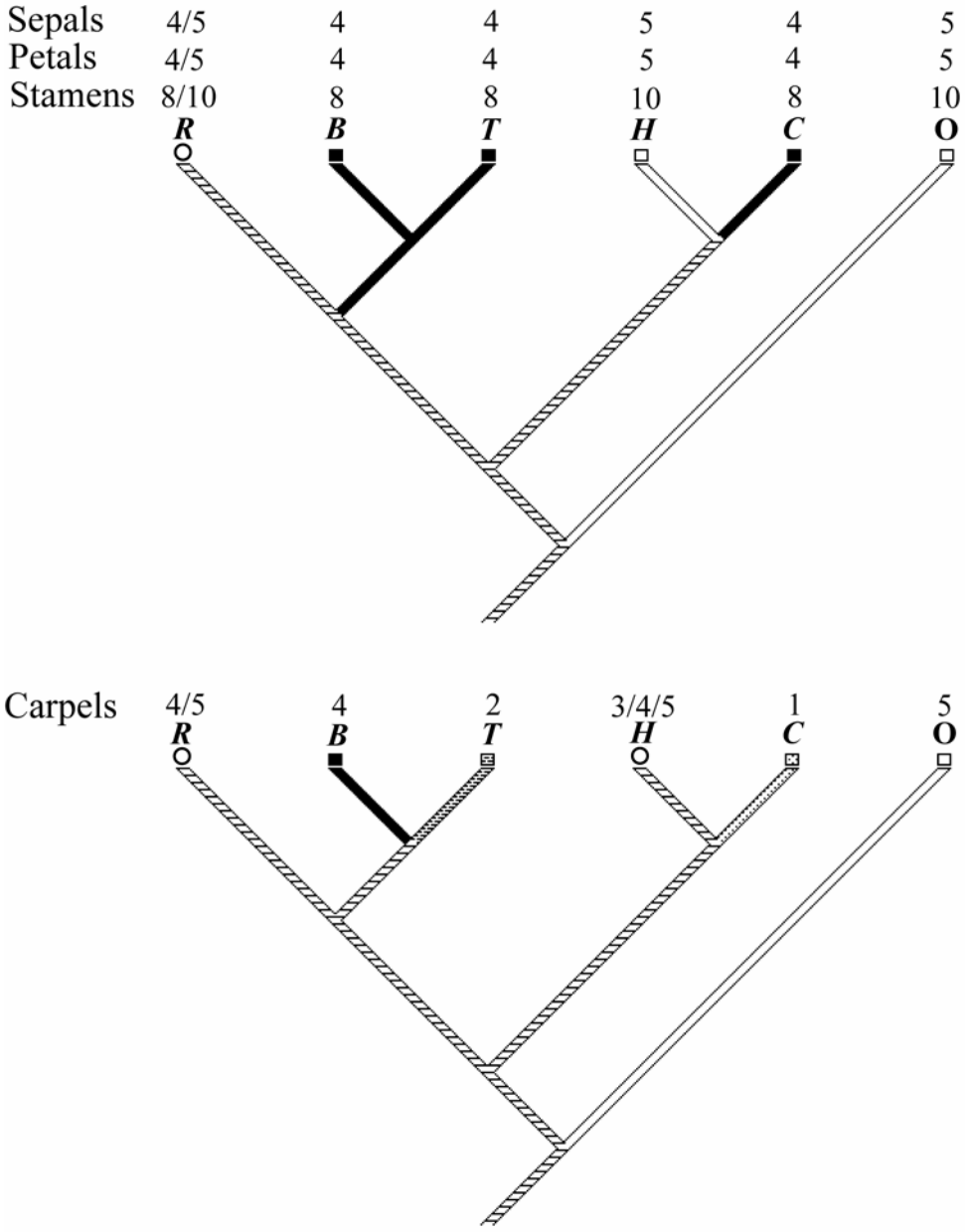

Sepals
Petals
Stamens
4/5
4/5
8/10
4
4
8
4
4
8
5
5
10
4
4
8
5
5
10
R
B
T
H
C
O
Carpels
4/5
4
2
3/4/5
1
5
R
B
T
H
C
O

Table 1. Engler's (1896, 1931) classification of Ruteae, with subsequent modifications by Townsend (1986) and Da Silva *et al*. (1988).

Engler (1896, 1931) Morphology	**Townsend (1986)** Morphology	**Da Silva *et al.* (1988)** Phytochemistry
Tribe Ruteae		
Subtribe Rutinae		***Ruta*-tribe**
Boenninghausenia	—	*Boenninghausenia*
Thamnosma	*Thamnosma*	*Thamnosma*
Cneoridium	—	*Cneoridium*
Ruta		
Subgenus *Euruta*	*Ruta*	*Ruta*
Subgenus *Haplophyllum*	*Haplophyllum*	*Haplophyllum*
Psilopeganum	—	—
Subtribe Dictamninae		***Dictamnus*-tribe**
Dictamnus	—	*Dictamnus*

Taxa not treated by the authors are indicated with a dash.

Table 2. Seven morphological characters, with their respective states, used by Engler (1896, 1931) to discriminate among the genera of tribe Ruteae.

	Ruta	***Thamnosma***	***Cneoridium***	***Boenninghausenia***	***Psilopeganum***	***Dictamnus***
Flower merism	4-merous or 5-merous	4-merous	4-merous	4-merous	4-merous	5-merous
Seed shape	angled dorsally	almost reniform	spherical	reniform	reniform	spherical
Number of ovules per locule	6 to 12	5 to 6	2	6 to 8	5 to 6	3
Number of carpels	4 or 5	2	1	4	2	5
Nectary disk shape	cushion-shaped	variable	ring-shaped	cup-shaped	small and narrow at the end	ring-shaped
Stigma shape	simple	capitate	almost spherical	simple	capitate	simple
Filament shape	broader at base	cuneate	cuneate	filiform	n.a.	club-shaped with protruding glands

The states of these characters are either overlapping or suggest different sister-group relationships (see text). "n.a." = not available.

Table 3. Results of S-H tests on two topological constraints.

Topologies	**No. of MP trees**	**MP length**	**-ln L scores**[a]	**-ln L differences**[b]	**P values**[c]
Unconstrained	6	2092	**17570.505**	-	-
Constraint (i): *Dictamnus* within Ruteae	94	2388	17626.010	55.506	0.080
Constraint (ii): *Haplophyllum* sister to *Ruta*	940	2455	17784.520	214.025	**<0.001**

[a]Highest likelihood scores assigned to MP trees; the highest likelihood score of the unconstrained trees (in boldface) was used for comparisons with the likelihood scores of all constrained trees. [b]Differences between likelihood scores of constrained and uncontraiend trees. Only the differences between the unconstrained tree with highest -ln L score and the constrained trees with highest -ln L scores are shown. [c]P values associated with differences between likelihood scores (significant values in boldface).

Table 4. Character and tree diagnostics, substitution models, and parameters estimated by ModelTest for the five data partitions.

	***mat*K**	***rpl*16**	***trn*L-*trn*F**	**Combined without gap coding**	**Combined with gap coding**
Aligned length	1564	1321	709	3594	3798
Parsimony informative ntps (% of aligned ntps)	429 (27.4)	294 (22.3)	125 (17.6)	848 (23.6)	957 (25.2)
No. of trees	600	803	70	6	16
No. of steps	1076	723	274	2092	2367
CI (CI ex)	0.744 (0.682)	0.765 (0.702)	0.803 (0.758)	0.752 (0.691)	0.751 (0.682)
RI	0.950	0.947	0.960	0.948	0.947
Model selected	TVM+G	TVM+G	TVM+G	TVM+G	TVM+G & Binary model
-lnL	8589.0039	6065.5581	2608.8774	17638.0703	-
Freq. [A]	0.2840	0.4054	0.4025	0.3525	-
Freq. [C]	0.1722	0.1450	0.1564	0.1575	-
Freq. [G]	0.1690	0.1522	0.1651	0.1658	-
Freq. [T]	0.3749	0.2974	0.2759	0.3242	-
R.r.o.s. [A-C]	1.1585	1.4023	0.5307	1.1031	-
R.r.o.s. [A-G]	1.5442	1.3557	0.8744	1.2461	-
R.r.o.s. [A-T]	0.2248	0.2041	0.0938	0.1931	-
R.r.o.s. [C-G]	0.8399	0.7504	0.6037	0.7632	-
R.r.o.s. [C-T]	1.5442	1.3557	0.8744	1.4942	-
R.r.o.s. [G-T]	1	1	1	1	-
I	0	0	0	0	-
Γ	0.8744	0.9871	0.5913	0.8191	-

For the "combined with gap coding" partition parameter values are not present because they were re-estimated by MrBayes. Ntps = nucleotide positions, CI = consistency index, CI ex = consistency index excluding parsimony uninformative characters, RI = retention index, Freq. = frequency, R.r.o.s. = relative rate of substitution, I = proportion of invariable sites, Γ = gamma distribution shape parameter.

Table 5. Character diagnostics and trees resulting from the analysis of the molecular, morphological, phytochemical, and global datasets.

Dataset	No. of characters	Parsimony informative characters	No. of optimal trees	No. of steps	CI (CI ex)	RI
Molecular	3176	179	1	673	0.924 (0.801)	0.791
Morphological	16	12	1	41	0.878 (0.828)	0.667
Phytochemical	31	18	1	38	0.816 (0.667)	0.667
Global	3223	209	1	769	0.899 (0.760)	0.721

CI = consistency index, CI ex = consistency index excluding parsimony uninformative characters, RI = retention index.

Appendix 1. Taxa used in this study, source, voucher information, and GenBank accession numbers for the three cpDNA regions studied. Voucher specimens are deposited in the following herbaria: Z = University of Zürich; K = Royal Botanic Gardens, Kew; MO = Missouri Botanical Garden; MA = Real Jardin Botanico, Madrid; FUMH = Ferdowsi University, Mashhad, Iran; GDA = Universidad de Granada; BR = Jardin Botanique National de Belgique. Exemplar accessions chosen to build the six-taxon molecular dataset are indicated with an asterix.

Species name	**Source**	**Voucher**	***mat*K**	***rpl*16**	***trn*L-*trn*F**
RUTACEAE					
Ruta graveolens L.*	Switzerland: Zürich Botanic Gardens; living collection	Cult. 19963642	EF489055	EF489129	EF489203
Ruta graveolens L.	France: Aveyron, Gages	Renaux 12 (Z)	EF489056	EF489130	EF489204
Ruta graveolens L.	Slovenia	J.Rothlisberger 6/7/99 (Z)	EF489057	EF489131	EF489205
Ruta chalepensis L.	Italy: Rio di Pula, Villa San Pietro, Calgliari, Sardinia	Bacchetta 50 (Z)	EF489044	EF489118	EF489192
Ruta chalepensis L.	France: La Bastide, Aude	Renaux 74 (Z)	EF489047	EF489121	EF489195
Ruta chalepensis L.	France: CNRS-Cargese, Corsica	Renaux 78 (Z)	EF489045	EF489119	EF489193
Ruta chalepensis L.	Greece: Pera Rachidhi, Katapola, Amorgos	Bacchetta & Brullo 117 (Z)	EF489046	EF489120	EF489194
Ruta chalepensis L.	Italy: Gutturu Cardaxius, Buggerru, Cagliari, Sardinia	Bachetta, Demurtas, and Pontecorvo 129 (Z)	EF489048	EF489122	EF489196
Ruta chalepensis L.	Italy: Caltabellotta, Sicily	G.Salvo 1 (Z)	EF489049	EF489123	EF489197
Ruta corsica DC.	Italy: Broncu Spina, Fonni, Nuoro, Sardinia	Bacchetta 162 (Z)	EF489050	EF489124	EF489198
Ruta corsica DC.	France: Col deVergio, Corsica	Renaux 185 (Z)	EF489052	EF489126	EF489200
Ruta corsica DC.	France: Asco, Corsica	Renaux 198 (Z)	EF489053	EF489127	EF489201
Ruta corsica DC.	Italy: Broncu Orisa, Villagrande Strisaili, Nuoro, Sardinia	Bacchetta & Carta 206 (Z)	EF489051	EF489125	EF489199
Ruta corsica DC.	France: Haut Ascò, Ascò, Bastia, Corsica	Adamo, Bacchetta, Carai, Iiriti, and Pontecorvo 226 (Z)	EF489054	EF489128	EF489202
Ruta angustifolia Pers.	France: Carriere de Scudo1, Corsica	Renaux 237 (Z)	EF489059	EF489133	EF489207

Ruta angustifolia Pers.	France : Gard, St-Hippolyte du Fort	Renaux 264 (Z)	EF489058	EF489132	EF489206
Ruta angustifolia Pers.	Spain: Barranc de Albranca, Ferries, Menorca	Bacchetta 283 (Z)	EF489061	EF489135	EF489209
Ruta angustifolia Pers.	Spain: Alzinar d'Alforì, Ciutadella, Menorca	Bacchetta, Fraga, and Vacca 311 (Z)	EF489060	EF489134	EF489208
Ruta montana (L.) L.	France: La Gardiole	Renaux 349 (Z)	EF489062	EF489136	EF489210
Ruta montana (L.) L.	Marocco: Atlas	Lewalle J. 9612 (Z)	EF489063	EF489137	EF489211
Ruta pinnata L.f.	Spain: Tenerife	MA 655741	EF489065	EF489139	EF489213
Ruta pinnata L.f.	Spain: Göimar, Tenerife, Canary Is.	Alfredo Amador 1	EF489066	EF489140	EF489214
Ruta oreojasme Webb & Berth.	Spain: Bco Tirajuana, Gran Canaria, Canary Is.	Alfredo Amador 3	EF489069	EF489143	EF489217
Ruta microcarpa Svent.	Spain: Parque Nacional de Garajonay, La Gomera, Canary Is.	Angel Fernandez 1	EF489067	EF489141	EF489215
Ruta microcarpa Svent.	Spain: La Gomera, Canary Is.	P.Vargas 20/10/05	EF489068	EF489142	EF489216
Boenninghausenia albiflora Reichb. ex Meissner*	Japan	Chase 22071 (K)	EF489070	EF489144	EF489218
Dictamnus albus L.	Switzerland: Zürich Botanic Gardens; living collection	Cult. 19964051	EF489110	EF489184	EF489258
Dictamnus albus L.	Bosnia-Herzegovina	M. Bandle & A. Lenherr 22483 (ZT)	EF489109	EF489183	EF489257
Cneoridium dumosum Hook.f.*	USA: Oak Crest Park, California	Alexander Kocyan 154 (Z)	EF489108	EF489182	EF489256
Thamnosma socotrana Balf.f.*	Yemen: Haghier, Socotra	Mike Thiv 3176 (Z)	EF489073	EF489147	EF489221
Thamnosma montana Torr. & Frem.	USA: Pinyon, Santa Rosa Mts., San Jacinto	Mike Thiv 4252 (Z)	EF489072	EF489146	EF489220
Thamnosma hirschii Schweinf.	Yemen: Hadibu, Socotra	Mike Thiv 3187 (Z)	EF489071	EF489145	EF489219
Thamnosma pailensis M.C.Johnst.	Mexico: Mpio. General Cepeda, Coahuila	Mike Thiv 369 (Z)	EF489074	EF489148	EF489222

Thamnosma texanum (A.Gray) Torrey	USA	M.Merello & D.Kruger 2643 (MO)	EF489075	EF489149	EF489223
Haplophyllum laeviusculum C.Towns.*	Iran	Zeltner 1b 5/5/8 (Z)	EF489083	EF489157	EF489231
Haplophyllum lissonotum C.Towns.	Iran	Zeltner 2b 5/5/15 (Z)	EF489084	EF489158	EF489232
Haplophyllum virgatum Spach	Iran	Zeltner 2 5/5/11 (Z)	EF489088	EF489162	EF489236
Haplophyllum affine (Aitch. & Hemsl.) Korovin	Iran	Zeltner 1 5/5/10 (Z)	EF489080	EF489154	EF489228
Haplophyllum bungei Trautv.	Iran	Zeltner 2 5/5/23 (Z)	EF489091	EF489165	EF489239
Haplophyllum pilosum Stschegleev ex Turcz.	Iran	F.Ghahremaninejad 1488 (Z)	EF489092	EF489166	EF489240
Haplophyllum versicolor Fisch. & Mey.	Iran	FUMH 23840	EF489090	EF489164	EF489238
Haplophyllum furfuraceum Bunge ex Boiss.	Iran	FUMH 14508	EF489093	EF489167	EF489241
Haplophyllum acutifolium (DC.) G.Don.	Iran	F.Ghahremaninejad 1469 (Z)	EF489076	EF489150	EF489224
Haplophyllum canaliculatum Boiss.	Iran	F.Ghahremaninejad 1454 (Z)	EF489077	EF489151	EF489225
Haplophyllum stapfianum Hand.-Mazz.	Iran	F.Ghahremaninejad 1426 (Z)	EF489078	EF489152	EF489226
Haplophyllum obtusifolium (Ledeb.) Ledeb.	Iran	FUMH 16876	EF489098	EF489172	EF489246
Haplophyllum tuberculatum (Forssk.)Adr.Juss.	Iran	F.Ghahremaninejad 1437 (Z)	EF489087	EF489161	EF489235
Haplophyllum robustum Bunge	Iran	FUMH 17725	EF489089	EF489163	EF489237
Haplophyllum glaberrimum Bunge ex Boiss.	Iran	FUMH 34555	EF489082	EF489156	EF489230
Haplophyllum suaveolens (DC.) G.Don.	Macedonia	MA 692105	EF489086	EF489160	EF489234

Haplophyllum coronatum Griseb.	Macedonia	MA 353234	EF489081	EF489155	EF489229
Haplophyllum patavinum (L.) G.Don	Bosnia-Herzegovina	MA 353204	EF489085	EF489159	EF489233
Haplophyllum villosum (M.Bieb.)G.Don	Azerbaijan	MA 417870	EF489096	EF489170	EF489244
Haplophyllum buxbaumii (Poir.) G.Don	Tunisia	MA 557457	EF489095	EF489169	EF489243
Haplophyllum linifolium (L.) G.Don	Spain: Almeria	MA 684635	EF489079	EF489153	EF489227
Haplophyllum bastentanum Navarro, Suarez-Santiago, & Blanca	Spain	GDA 47502	EF489097	EF489171	EF489245
Haplophyllum latifolium Kar. & Kir.	Uzbekistan	MA 642325	EF489094	EF489168	EF489242
Haplophyllum dauricum (L.) G.Don	Mongolia	Oyumaa 1 (Z)	EF489099	EF489173	EF489247
Calodendrum capense (L.f.) Thunb.	Switzerland: Zürich Botanic Gardens; living collection	Cult. 20010674	EF489102	EF489176	EF489250
Evodia hupehensis Dode.	Switzerland: Zürich Botanic Gardens; living collection	Cult. 19820014	EF489105	EF489179	EF489253
Zanthoxylum simulans Hance	Switzerland: Zürich Botanic Gardens; living collection	Cult. 19963815	EF489100	EF489174	EF489248
Zanthoxylum americanum P.Mill.	Spain: Real Jardin Botanico of Madrid; living collection	Gonzalo Nieto Feliner 4/22/05	EF489101	EF489175	EF489249
Choisya ternata Kunth*	Switzerland: Zürich Botanic Gardens; living collection	Cult. 19963167	EF489104	EF489178	EF489252
Melicope sp.	New Caledonia	J. Munzinger & G. McPherson 785 (MO)	EF489107	EF489181	EF489255
Skimmia japonica Thunb.	Switzerland: Zürich Botanic Gardens; living collection	Cult. 19651137	EF489103	EF489177	EF489251

Orixa japonica Thunb.	Belgium: National Botanic Garden of Belgium; living collection	Cult. 19850295-..	EF489106	EF489180	EF489254
SIMAROUBACEAE					
Ailanthus altissima (Mill.) Swingle	Tibet	Chase 16978 (K)	EF489111	EF489185	EF489259
Ailanthus giraldii Dode		Chase 16984 (K)	EF489112	EF489186	EF489260
Simarouba glauca DC.		Chase 124 (K)	EF489113	EF489187	EF489261
MELIACEAE					
Swietenia macrophylla King		Chase 250 (K)	EF489114	EF489188	EF489262
Azadirachta indica A.Juss.	Switzerland: Zürich Botanic Gardens; living collection	Cult. 19940741	EF489115	EF489189	EF489263
Cipadessa baccifera (Roth) Miq.	Switzerland: Zürich Botanic Gardens; living collection	Cult. 20011380	EF489116	EF489190	EF489264
Melia azedarach L.	Switzerland: Zürich Botanic Gardens; living collection	Cult. 20012236	EF489117	EF489191	EF489265

Appendix 2. List of morphological characters, and corresponding character-states, that vary among *Ruta*, *Boennighausenia*, *Thamnosma*, *Haplophyllum*, and *Cneoridium*. Characters were scored from descriptions in Engler (1931) and Townsend (1986). In the matrix polymorphic states are in parentheses and missing states are indicated with a question mark.

Morphological characters

1 Number of petals: **0** = 5 petals; **1** = 4 petals
2 Number of sepals: **0** = 5 sepals; **1** = 4 sepals
3 Number of stamens: **0** = 10 stamens; **1** = 8 stamens
4 Number of carpels: **1** = 1 carpel, **2** = 2 carpels, **3** = 3 carpels, **4** = 4 carpels, **5** = 5 carpels
5 Petal shape: **0** = obovate; **1** = oblong; **2** = ovate
6 Sepal shape: **0** = ovate; **1** = deltoid; **2** = linear
7 Shape of nectary disk: **0** = cushion-shaped; **1** = cup-shaped; **2** = ring-shaped
8 Filament shape: **0** = lanceolate; **1** = broader at the base; **2** = cuneate; **3** = filiform
9 Style: **0** = short and thick; **1** = thin and long
10 Petal margins: **0** = entire; **1** = dentate or fimbriate
11 Anther opening: **0** = lateral; **1** = slightly introrse; **2** = introrse
12 Stigma shape: **0** = almost spherical (a bit lobed); **1** = capitate; **2** = simple
13 Anther shape: **0** = globular; **1** = elongate; **2** = ovate
14 Seed shape: **0** = spherical; **1** = reniform; **2** = angled dorsally
15 Embryo: **0** = curved; **1** = not curved
16 Gynophore: **0** = yes; **1** = no

Morphological matrix

Character	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Ruta	(01)	(01)	(01)	(45)	0	(01)	0	1	0	1
Boenninghausenia	1	1	1	4	0	0	1	3	1	0
Thamnosma	1	1	1	2	2	0	(01)	2	1	0
Haplophyllum	0	0	0	(345)	(12)	(012)	0	1	0	0
Cneoridium	1	1	1	1	2	0	2	2	0	0
Choisya	0	0	0	5	0	0	?	0	0	0

Character	**11**	**12**	**13**	**14**	**15**	**16**
Ruta	2	2	1	2	1	1
Boenninghausenia	1	2	2	1	0	0
Thamnosma	1	1	2	(012)	0	0
Haplophyllum	2	0	1	1	1	1
Cneoridium	2	0	0	0	0	1
Choisya	0	0	0	0	0	0

Appendix 3. List of phytochemical characters, and corresponding character-states, that vary among *Ruta*, *Boennighausenia*, *Thamnosma*, *Haplophyllum*, and *Cneoridium*. Characters were scored from data available in Da Silva *et al*. (1988).

Phytochemical characters

17 2-arylquinazolines of type C1
18 2-arylquinolines/ones of type D1
19 2-arylquinolines/ones of type D1.1
20 2-alkylarylquinolines/ones of type E1
21 2-alkylquinolines/ones of type F1
22 2-quinolones of type G1
23 2-quinolones of type G1.1
24 2-quinolones of type G1.1.1
25 2-quinolones of type G1.1.2
26 2-quinolones of type G1.1.2.1
27 2-quinolones of type G1.1.2.2
28 2-quinolones of type G1.1.2.2.1
29 2-quinolones of type G1.1.1.2
30 2-quinolones of type G1.1.1.3
31 2-quinolones of type G1.2
32 2-quinolones of type G1.8
33 2-quinolones of type G1.9
34 acridones of type H1
35 acridones of type H1.1
36 acridones of type H1.1.1
37 acridones of type H1.2
38 coumarins of type O1.1
39 coumarins of type O1.1.1
40 coumarins of type O1.2
41 coumarins of type O1.5
42 coumarins of type O1.5.1
43 coumarins of type O1.6
44 flavonoids of type P1.2.1
45 lignans of type Q1
46 lignans of type Q1.1
47 lignans of type Q

Phytochemical matrix (1 = present, 0 = absent)

Character	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**
Ruta	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
Boenninghausenia	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Thamnosma	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Cneoridium	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Haplophyllum	0	1	1	0	1	1	1	1	1	0	0	1	1	1	1
Choisya	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0

Character	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**
Ruta	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1
Boenninghausenia	0	0	1	1	0	1	1	1	1	1	0	0	1	0	1
Thamnosma	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0
Cneoridium	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1
Haplophyllum	1	1	0	0	0	0	1	0	1	0	0	1	1	1	1
Choisya	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

Character	**47**
Ruta	1
Boenninghausenia	0
Thamnosma	0
Cneoridium	0
Haplophyllum	1
Choisya	0

Chapter II

Tracing the temporal and spatial origins of island endemics in the Mediterranean region: a case study from the citrus family (*Ruta* L., Rutaceae)

Gabriele Salvo[1*], Simon Y. W. Ho[2], Gideon Rosenbaum[3], Richard Ree[4], Elena Conti[1]

[1]*Institute of Systematic Botany, University of Zürich, CH-8008 Zürich, Switzerland*.

[2]*Centre for Macroevolution and Macroecology, Research School of Biology, Australian National University, Canberra ACT 0200, Australia*.

[3]*School of Earth Sciences, University of Queensland, St. Lucia, Queensland 4072, Australia*.

[4]*Botany Department, Field Museum of Natural History, 60605 Chicago, USA*.

Published in:

Systematic Biology doi: 10.1093/sysbio/syq046

Abstract

Understanding the origin of island endemics is a central task of historical biogeography. Recent methodological advances provide a rigorous framework to determine the relative contribution of different biogeographic processes (e.g., vicariance, land migration, long distance dispersal) to the origin of island endemics. With its complex but well-known history of microplate movements and climatic oscillations, the Mediterranean region (including the Mediterranean basin and Macaronesia) provides the geographic backdrop for the diversification of *Ruta* L., the type genus of Rutaceae (citrus family). Phylogenetic, molecular dating, and ancestral range reconstruction analyses were carried out to investigate the extent to which past geological connections and climatic history of the Mediterranean region explain the current distribution of species in *Ruta*, with emphasis on its island endemics. The analyses showed that *Ruta* invaded the region from the north well before the onset of the Mediterranean climate, and diversified *in situ* as the climate became Mediterranean. The continental fragment island endemics of the genus originated via processes of land migration/vicariance driven by connections/disconnections between microplates, whereas the oceanic island endemics were the product of a single colonization event from the mainland followed by *in situ* diversification. This study emphasizes the need for an integrative, hypothesis-based approach to historical biogeography and stresses the importance of temporary land connections and colonization opportunity in the biotic assembly of continental fragment and oceanic islands, respectively.

Key words

Historical biogeography, Mediterranean region, oceanic islands, continental fragment islands, *Ruta*, geologic history, palaeo-climate.

Introduction

Historical biogeography has evolved from a mostly pattern-oriented exercise to an integrative, model-based discipline. The past ten years have witnessed a resurgence of biogeographic studies that strive to integrate information from phylogenies, fossils, molecular dating, the geologic record, and palaeo-climatic reconstructions. In particular, the following have become instrumental in setting up sound historical biogeography hypotheses: i) the establishment of a temporal framework (Hunn and Upchurch 2001; Donoghue and Moore 2003); ii) the *a priori* incorporation of explicit models pertaining to relevant biological and earth processes into biogeographic analysis (Ree et al. 2005; Sanmartín et al. 2008; Ree and Sanmartín 2009); and iii) the shift in focus from the vicariance-versus-dispersal dualism to a more quantitative assessment of the contribution of different historical processes, including geodispersal and extinction, to current patterns of distribution (Crisci et al. 2003; Lieberman 2003; Lomolino et al. 2006; Upchurch 2008).

The integration of different kinds of data/models into biogeographic analysis introduces more variables into the reconstruction of biogeographic scenarios, often producing an increase in uncertainty with respect to the inferences being made (Conti et al. 2004; Heads 2005; Rutschmann et al. 2007; Clark et al. 2008; Nylander et al. 2008; Ho and Phillips 2009). A crucial step in minimizing uncertainty, and a prerequisite for sound analyses in historical biogeography, is the choice of a group of organisms for which: i) a robust phylogenetic framework is present; ii) reliable fossils for molecular clock calibration are available; iii) obvious traits facilitating long distance dispersal (hereafter, LDD) are absent; and iv) the geology and palaeo-climate of the area housing them are well understood. The focal area (the Mediterranean region) and taxon (*Ruta* L.) selected for the present study satisfy these requirements.

With its complex but well-known history of micro-plate movements and climatic oscillations, the Mediterranean region provides the geographic backdrop for the diversification of *Ruta*, the type genus of Rutaceae (citrus family). As currently circumscribed, *Ruta* includes nine species of perennial herbs: *R. angustifolia*, *R. chalepensis*, *R. montana*, and *R. graveolens* exhibit a circum-Mediterranean distribution; *R. corsica* is endemic to Corsica and *R. lamarmorae* to Sardinia; *R. pinnata*, *R. oreojasme*, and *R. microcarpa* are endemic to the Canary Islands (Townsend 1968; Bramwell and Bramwell 2001; Bacchetta et al. 2006; Fig. 1). The limits of the geographic distribution of *Ruta* broadly correspond to the limits of the Mediterranean region, and the genus often occurs in association with elements characteristic of the Mediterranean vegetation (for example, *Ulex*, *Quercus*, *Pistacia*; Bonet 1992). Thus, the biogeographic history of *Ruta* has broad implications for the

assembly of the Mediterranean flora. A recent phylogenetic analysis of *Ruta* and closely related genera ascertained the monophyly of the genus (Salvo et al. 2008). Additionally, the palaeontological record of Rutaceae is quite rich (Gregor 1989) and *Ruta* produces capsules and seeds without any obvious adaptations for wind dispersal (Engler 1896, 1931), two factors that contribute to making *Ruta* an ideal genus for a biogeographic analysis.

The evolution of diversity in *Ruta* has not been examined in a temporal framework, nor has its current distribution been analyzed in the context of the geologic history and palaeo-climate of the Mediterranean basin. Most importantly, the occurrence of five out of nine species endemic to either continental fragment (Corsica and Sardinia) or oceanic (the Canarian archipelago) islands offers a unique opportunity for comparing modes of biogeographic evolution in these two kinds of islands. The main difference between the biogeographic mechanisms that can be invoked to explain the origin of the five island endemics is that the colonization of the oceanic Canary Islands could have been achieved exclusively via LDD, whereas colonization of the continental fragment islands of Corsica and Sardinia could have been effected via both LDD and land migration (Emerson 2002; Cowie and Holland 2006; Lomolino et al. 2006; Sanmartín et al. 2008). Below we summarize the geologic history, main palaeo-climatologic features, and hypotheses on the origin of the flora of the focal area, with special emphasis on Corsica, Sardinia, and the Canary Islands.

With its 22,500 species and about 13,000 endemics, the Mediterranean region – comprising the Mediterranean basin and the Macaronesian islands (Quézel 1985) – has been recognized as a hotspot of biodiversity (Medail and Quézel 1997; Myers et al. 2000). A large component of the Mediterranean basin formed during the Cenozoic, simultaneously with the ongoing convergence of Africa with respect to Europe (Dercourt et al. 1986; Dewey et al. 1989; Rosenbaum et al. 2002b). The present land-locked configuration of the basin resulted from the collision of the Arabian plate with stable Eurasia in the middle Miocene, which led to the closure of the connection between the Tethys Sea and the Indian Ocean (Krijgsman 2002; Garfunkel 2004). In the early Miocene (23-16 million years ago, MA) and beyond, the basin experienced subtropical conditions, with little seasonal change in temperature and relatively high levels of summer rainfall (Thompson 2005). In the middle Miocene (16-14 MA), seasonal contrasts in the temperature regime started to develop, leading to the establishment of the current Mediterranean climatic rhythm (summer drought) in the Pliocene (3-2 MA; Suc 1984; Thompson 2005). The Pleistocene (1.8-0.01 MA) witnessed the alternation of warmer and colder conditions during the glacial cycles, while increased aridity and gradual warming marked the Holocene (0.01 MA-present).

Both the beginning of a trend towards increasing aridification in the Mediterranean region (9-8 MA) and the onset of the Mediterranean climate (around 3-2 MA) are believed to have had a big impact on the composition and structure of the Mediterranean flora (Suc 1984; Ivanov et al. 2002; Thompson 2005; Van Dam 2006; Lo Presti and Oberprieler 2009). The former event led to the progressive replacement of the tropical elements of the Mediterranean flora (e.g., mangroves, Taxodiaceae) with sclerophyllous plant communities; the latter event caused coastal forests to disappear and xerophytic taxa (e.g., *Olea*, *Quercus ilex*-type, *Pistacia*) to expand (Suc 1984; Thompson 2005). Early workers temporally divided the floristic elements of the Mediterranean region into two main groups, depending on whether they were believed to have originated before or after the development of such climate (Pons and Quézel 1985; Thompson 2005).

A series of events likely to have profoundly influenced the biogeography of the Mediterranean basin concerns the geologic history of the continental fragment islands of Corsica and Sardinia (Palmer 1998; Caccone et al. 1994; Caccone and Sbordoni 2001; Ketmaier et al. 2003, 2006; Fochetti et al. 2004). Before the early Oligocene, these two islands were situated adjacent to current southern France, forming a continuous geologic entity (part of the so-called Hercynian massif) that was subsequently fragmented into microplates that dispersed throughout the western Mediterranean (Alvarez et al. 1974; Rosenbaum et al. 2002a). These tectonic fragments included the Tuscan archipelago (in Italy), the Balearic Islands, the internal parts of the Betic-Rif *cordillera* (in Spain and Morocco, respectively), the Kabylies (in Algeria), and the Calabro-Peloritan massif (in southern Italy). According to tectonic reconstructions, in the late Oligocene (30-28 MA) the Balearic-Kabylies microplate and the Corso-Sardinian-Calabro-Peloritan microplate separated from the eastern part of the proto-Iberian peninsula (Alvarez et al. 1974; Dewey et al. 1989; Rosenbaum et al. 2002a). At around 25 MA, the Balearic-Kabylies microplate started to rotate clockwise until the Balearic Islands reached their current position (around 21 MA) and then became detached from the Kabylies terrane, which continued to drift towards North Africa. At the same time the Corso-Sardinian-Calabro-Peloritan microplate moved eastward, with a counter-clockwise rotation of approximately 30°, until it collided with the western side of the Apulian microplate around 20-18 MA (Cherchi and Montadert 1982; Deino et al. 2001; Rosenbaum et al. 2002a; Speranza et al. 2002). These two microplates remained connected until the opening of the North Tyrrhenian Sea in the Tortonian (around 9 MA; Rosenbaum and Lister 2004a). The Calabro-Peloritan block finally split off from the Corso-Sardinian (C-S) microplate around 5 MA, reaching its current position as the southern-most tip of the Italian peninsula during the last 2 million years (MY; Rosenbaum et al. 2002a). During the

Messinian Salinity Crisis (MSC; 5.96-5.33 MA), an event characterized by the dramatic drying and salinity increase of the Mediterranean Sea (Robertson and Grasso 1995; Krijgsman 2002), the C-S microplate was most likely linked with the south-western Alps and present-day Tuscany (Italy) to the north and with North Africa to the south (Rögl and Steininger 1983; Rouchy and Saint-Martin 1992; Robertson and Grasso 1995; Gover et al. 2009).

The tectonic separation of Corsica from Sardinia began 15 MA and was complete by 9 MA (Alvarez 1972, 1976; Alvarez et al. 1974; Bonin et al. 1979; Orsini et al. 1980; Cherchi and Montadert 1982), although episodic contacts between these two islands persisted until very recently, in particular during the MSC (Rögl and Steininger 1983; Gover et al. 2009) and the Last Glacial Maximum (LGM; around 0.02 MA; Lambeck et al. 2004; Lambeck and Purcell 2005). At present, the two islands are situated in the middle of the western Mediterranean basin and are separated by a narrow (11 km) and shallow (less than 60 meters deep) water channel, the Strait of Bonifacio.

Corsica and Sardinia have been identified as one of the ten areas with the highest biodiversity in the Mediterranean (Médail and Quézel 1997), with about 4300 plant species, including around 340 listed endemics and subendemics (Arrigoni 1979; Gamisans and Jeanmonod 1993). A long tradition of floristic studies on the C-S flora has revealed affinities with the Pyrenees, Provençe, southern Spain, the Balearic Islands, Liguria, Calabria/Sicily, and North Africa (Moris 1837-59; Barbey 1885; Briquet 1910; Braun-Blanquet 1926; Contandriopoulos 1962; Bocquet et al. 1978; Jeanmonod and Gamisans 1987). Five main scenarios have been advanced to explain the origin of the C-S endemic flora: i) the split between the C-S microplate and the proto-Iberian peninsula (30-28 MA), causing vicariant speciation in taxa that inhabited both areas before the split (Gamisans 1975); ii) subsequent land connections between the C-S microplate and adjacent areas, which facilitated the entering of floristic elements into the two islands (Braun-Blanquet 1926; Contandriopoulos 1962); iii) the MSC (5.96-5.33 MA), which, through the mesh of land corridors present in the Mediterranean basin at that time, allowed a "veritable explosion in migrations" across the basin, affecting also Corsica and Sardinia (Bocquet et al. 1978; p. 277); iv) the Quaternary glacial cycles, which, due to the supposed connections between the C-S block and present-day Provençe/Liguria/Tuscany at that time, enhanced biotic exchange (Mariotti 1990); and v) LDD, first proposed by Engler (1879) to explain the origin of the Corsican flora. These scenarios, however, were based on floristic studies that lacked an explicit spatio-temporal framework, and sometimes used the distribution of organisms to infer past palaeo-geographical settings, rather than *vice versa*.

The Canary Islands, in the Atlantic Ocean, off the coast of Africa, are part of the Macaronesian phytogeographical region, which also includes the Azores, Madeira, the Salvage Islands, and the Cape Verde Islands (Sunding 1979). These islands were formed in an East-West progression, starting with Fuerteventura (around 20 MA) and ending with El Hierro (around 1 MA; Carracedo et al. 1998; Anguita and Hernán 2000). The Canaries have a Mediterranean climate, generally similar to that of Corsica and Sardinia (Whittaker and Fernandez-Palacios 2007), and are floristically the richest islands of Macaronesia, with around 680 endemic species (Reyes-Betancort et al. 2008). Interest in the phylogenetic origin of the Canarian biota has significantly increased in recent years (Juan et al. 2000; Carine et al. 2004; Kim et al. 2008; Sanmartín et al. 2008). From a temporal point of view, the endemic flora of the Canary Islands has been viewed either as the relict of a formerly widespread subtropical flora that covered southern Europe and North Africa during the Tertiary, or as a relatively younger flora (Emerson 2002; Carine 2005). In any case, since the Canary Islands are a volcanic archipelago and were thus never connected to the mainland, LDD represents the only possible mode of colonization of the archipelago (Cowie and Holland 2006; Sanmartín et al. 2008). Different LDD scenarios have been proposed for the colonization of these islands, invoking, at one extreme, a single LDD event to the older island, followed by progressive colonization of the younger islands in a stepping-stone fashion, or, at the other extreme, multiple LDD events in a sequence that is independent from the order of island emergence (Juan et al. 2000; Cowie and Holland 2006; Sanmartín et al. 2008).

By adopting an integrative approach, including phylogenetic, molecular dating, and ancestral range reconstruction analyses, in the context of the geologic history and palaeo-climate of the Mediterranean region, we ask: I) Was the origin and diversification of *Ruta* concomitant with the onset of the Mediterranean climate, and did its ancestor evolve *in situ* or in areas neighbouring the Mediterranean region? II) Does the fragmentation of the Hercynian massif explain the origin of the Corso-Sardinian lineage? III) Does the formation of the Strait of Bonifacio explain the divergence between the two Corsican and Sardinian endemic species, respectively? IV) Does the colonization of the Canary Islands by *Ruta* conform to a stepping-stone mode of island diversification, and was the origin of the three Canarian endemics concomitant with the onset of the Mediterranean climate?

Materials and Methods

Taxon sampling

All extant species of *Ruta* were sampled (Townsend 1968; Bramwell and Bramwell 2001; Bacchetta et al. 2006). The sampling of *Ruta*'s most closely-related taxa was guided by

careful examination of taxonomic (Engler 1896, 1931), phytochemical (Da Silva et al. 1988), and phylogenetic studies (Chase et al. 1999; Scott et al. 2000; Poon et al. 2007; Groppo et al. 2008; Salvo et al. 2008) on Rutaceae and included: *Thamnosma*, *Boenninghausenia*, *Haplophyllum*, *Cneoridium*, and representatives of subfamily Aurantioideae. Since available fossils have been linked with extant taxa that are distantly related to *Ruta* (see below), species representative of the phylogenetic diversity of Rutaceae, including the fossil-bearing taxa, were sampled according to the family-level phylogenetic analysis of Groppo et al. (2008). Where possible, two species per genus were included for the fossil-bearing taxa, in order to discriminate between their crown and stem nodes for fossil calibration purposes (see Magallón 2004). The final matrix contained 48 accessions, including 44 from Rutaceae and two each from Meliaceae and Simaroubaceae, which served as outgroups (Gadek et al. 1996; Muellner et al. 2007). Included material, voucher information, sources, and GenBank/EBI accession numbers are listed in online Appendix 1 (available from http://www.sysbio.oxfordjournals.org).

DNA sequences and phylogenetic analyses

Three chloroplast markers that provided sufficient resolution at our level of investigation and allowed unequivocal alignments were chosen: the coding region of the *mat*K gene, amplified using primers 1F and 1R (Sang et al. 1997); the *rpl*16 intron, amplified using primers F71 and R1516 (Baum et al. 1998); and the intergenic spacer between the *trn*L (UAA)3' exon and *trn*F (GAA) (from hereafter, *trn*L-*trn*F intergenic spacer), amplified with primers e and f (Taberlet et al. 1991). DNA extraction, PCR amplification and sequencing followed the methods described in Salvo et al. (2008). Sequences were edited and assembled using Sequencher 4.2™ software (Gene Codes Corp., Ann Arbor, Michigan, USA). Base positions were individually double-checked for agreement between the complementary strands. All sequences were visually aligned in MacClade 4.06 (Maddison and Maddison 2000). Regions of ambiguous alignment were excluded from the analysis. Gap positions were treated as missing data, unequivocally aligned gaps being coded as presence/absence of characters with the software GapCoder (Young and Healy 2003) and then added as binary characters to the data matrix.

Three data partitions were defined, corresponding to the three loci of the chloroplast genome examined in this study. We are aware of the important issue of partitioning in multi-locus molecular studies, and of the problems of under- and over-parameterization (Nylander et al. 2004; Brown and Lemmon 2007); however, exploratory analyses using different partitioning-schemes did not have a noticeable impact on the resulting chronograms. The individual partitions were initially analyzed separately to establish whether any strongly

supported (i.e., > 85 bootstrap percentage, BP) clades were incongruent among the respective trees. Since no such incongruence was detected, the sequences of the three loci were combined in a single dataset. The combined matrix, which is available at TreeBASE (http://purl.org/phylo/treebase/phylows/study/TB2:S10145), was then analyzed using maximum parsimony (MP), maximum likelihood (ML), and Bayesian Inference (BI). Parsimony analyses were conducted using PAUP*4.0b10 (Swofford 2001). All changes were treated as unordered and equally weighted. Tree search was performed using the following protocol: (i) a heuristic search was carried out with 1000 replicates of random taxon addition sequence and 10 trees held at each step, and tree bisection-reconnection branch swapping (TBR) on best trees only, with no more than 100 trees saved per replicate; (ii) the best trees found in (i) were then used as starting trees for a second heuristic search using TBR branch swapping until all swapping options were explored, and saving multiple trees. The STEEPEST DESCENT option was used in both (i) and (ii). Relative support for each node obtained by MP was assessed using bootstrap re-sampling (Felsenstein 1985). The following protocol was employed: heuristic search, 1000 bootstrap replicates, 100 random addition sequence replicates with three trees held at each step, TBR swapping with STEEPEST DESCENT and saving no more than ten trees per replicate.

Maximum-likelihood analyses were performed using RAxML (Stamatakis 2006). The GTR+G model of nucleotide substitution was used, with a separate model for each data partition. Support values for nodes in the phylogenetic tree were obtained by analysing 1000 pseudoreplicate data sets generated through bootstrap sampling from the original alignment.

Bayesian inference was performed in MRBAYES v3.1.2 (Hulsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003), after determining the model of evolution most suitable for each individual cpDNA region with the Akaike Information Criterion (Akaike 1974) in ModelTest 3.06 (Posada and Crandall 1998). The GTR+G model of nucleotide substitution was selected for all three regions. Two independent runs with four Monte Carlo Markov chains (MCMCs, one cold and three incrementally heated) run for 5×10^6 generations, with trees sampled every 1000^{th} generation, were performed. Each chain used a random tree as starting point and the default temperature parameter value of 0.2. The first 25000 sampled trees were discarded as “burn in” after checking for stability on the log-likelihood curves and after visual inspection of the split (clade) frequencies using the software AWTY (Wilgenbusch et al. 2004). The remaining trees were used to build a 50% majority rule consensus tree.

Selection of fossils for calibration

Rutaceae are well represented in the fossil floras of North America, Europe, and Asia, extending back into the latest Cretaceous. The most comprehensive revision of the fossil record of the family was carried out by Gregor (1989), who used morphological characters of the seeds' testa, inner anatomy, and surface in order to differentiate fossil taxa and determine affinities with extant genera. The oldest certain Rutaceae fossil is *Rutaspermum biornatum* (around 65 MA; Maastrichtian of Walbeck, Germany; Knobloch and Mai 1986), followed by fossils belonging to the extant genera *Zanthoxylum*, *Euodia*, *Acronychia*, *Toddalia*, and *Fagaropsis* (Gregor 1989). Fossils of *Ruta* have not been found. Four fossils accompanied by detailed morphological descriptions, dates associated with the geologic interval of their collection locality, and well-supported affinities with modern taxa were selected: *Euodia lignita* Tiffney (Tiffney 1980, 1994), *Toddalia excavata* (Chandler) Gregor (Gregor 1979), *Ptelea enervosa* H.V. Smith (Smith 1938), and *Skimmia tortonica* Palamarev and Usunova (Palamarev and Usunova 1970). Detailed information on these fossils is presented in online Appendix 2.

Molecular dating analyses

In order to investigate the degree of substitution rate variation among lineages, a likelihood-ratio test (LRT) was performed using PAUP*. First, the best model of evolution for the combined matrix without gap coding was selected using ModelTest. Second, the parameters describing the selected model were used to compute the likelihood score of the Bayesian 95% majority rule consensus tree with and without enforcing substitution rate constancy. Since the LRT rejected rate constancy, molecular dating analyses were carried out using two Bayesian methods that allow for variation of substitution rate among tree branches. The first method, which assumes rate autocorrelation between neighbouring branches, was implemented in MULTIDIVTIME (Thorne et al. 1998; Kishino et al. 2001; Thorne and Kishino 2002). The second method, where different models of among-lineage rate variation can be used, and priors on calibrations can be modelled with parametric distributions, was implemented in BEAST v1.4.8 (Drummond et al. 2006; Drummond and Rambaut 2007).

MULTIDIVTIME analyses were carried out using the following protocol. After having estimated branch lengths and their variance-covariance matrix with the BASEML and ESTBRANCHES programmes, contained in the PAML 4 (Yang 2007) and Multidistribute v. 9/25/03 (Thorne et al. 1998; Thorne and Kishino 2002) packages, respectively, we performed MCMC searches in MULTIDIVTIME. Four prior distributions were specified in units of 10 MY: the mean of the ingroup's age (rttm) was set to 6.5, because the oldest reliable fossil of Rutaceae was dated to 65 MA (Knobloch and Mai 1986); the standard deviation of the

ingroup's age (rttmsd) was set to 6.0, because an ingroup age older than the appearance of tricolpate pollen in the fossil record (around 125 MA; Sanderson and Doyle 2001; Anderson et al. 2005) was deemed unrealistic; the mean (rtrate) and standard deviation (rtratesd) of the rate at the root node were set to 0.0096 substitutions/site/10 MA, by dividing the median of the distances between the ingroup root and the tips by rttm; the mean (brownmean) and standard deviation (brownsd) of the Brownian motion parameter nu were set to 0.23 units, so that brownmean × rttm = 1.5. Finally, the age of a larger clade to which the group belongs (bigtime) was set to 12.5, referring again to the emergence of tricolpate pollen in the fossil record. Internal calibrations were set up using the four fossils mentioned above: the upper (younger) bound of the geologic interval in which each fossil was found represented the minimum age constraint. An age of 1.597 was assigned to the stem node of *Euodia*, an age of 3.72 to the node subtending *Toddalia*, an age of 1.1608 to the stem node of *Ptelea*, and an age of 0.7246 to the node subtending *Skimmia*. The Bayesian 50% majority rule consensus tree obtained with MRBAYES was used as the input tree. The MCMC was run for 10^6 generations and sampled every 100^{th} generation, after an initial burn-in period of 10^5 generations. Two separate runs were conducted to check for convergence on similar posterior distributions. To assess the effect of choice of priors on resulting age estimates, we performed sensitivity analyses by selecting different root prior distributions and by repeating all analyses with a fully-resolved input tree, the maximum a posteriori (MAP) tree inferred with MRBAYES (online Appendix 3).

BEAST simultaneously estimates phylogenetic relationships and nodal ages. The age estimates are calibrated through the specification of prior age distributions for certain nodes in the tree, guided by independent information from the fossil record. Since the oldest fossilized representative of a clade corresponds to the minimum age of that clade, parametric distributions can be modelled around fossil calibration points, in which the probability of a clade being younger than its oldest fossil drops immediately to zero, but the probability of it being older than its oldest fossil decays more gradually (Ho and Phillips 2009). Calibrations using the four fossils described above were modelled with a log-normal distribution, where 95% of the prior weight fell within the geologic interval in which each fossil was discovered (online Appendix 2). For *Euodia lignita* (33.9 to 15.97 MA) the parameters of the log-normal calibration prior were: hard minimum bound 15.97, mean 2.1933, and standard deviation 0.4214. For *Toddalia excavata* (40.4 to 37.2 MA) the parameters were: hard minimum bound 37.2, mean 0.47, and standard deviation 0.4215. For *Ptelea enervosa* (15.97 to 11.608 MA): hard minimum bound 11.608, mean 0.7798, and standard deviation 0.4214. And for *Skimmia tortonica* (11.608 to 7.246 MA): hard minimum bound 7.246, mean 0.7798, and standard

deviation 0.4214. The prior age of the root was also modelled with a log-normal distribution, with the following parameters: hard minimum bound 65, mean 2.3, and standard deviation 0.7. The mean of this distribution corresponded to the oldest, reliable Rutaceae fossil, and the 1% lower tail to the appearance of tricolpate pollen in the fossil record (see above). After having selected a GTR substitution model with 4 gamma categories, an uncorrelated, lognormal, relaxed-clock model and a Yule prior on the tree, five independent runs of 10^7 generations each, sampling every 1000^{th} generation, with a burn-in of 10^6 generations each, were conducted. The post-burnin trees from each run were combined and a maximum-clade-credibility tree was computed. Similarly to MULTIDIVTIME, sensitivity analyses were carried out to assess the impact of different root priors and prior distributions for fossil calibrations on the posterior distribution of age estimates (online Appendix 3).

Since the node to which the fossil *Ptelea enervosa* was attached received weak support, in terms of both PP and BP values (Fig. 2), MULTIDIVTIME and BEAST analyses were repeated omitting this fossil calibration.

Ancestral range reconstruction analyses

To infer ancestral areas and geographic speciation scenarios for *Ruta* and its sister group, we implemented a dispersal-extinction-cladogenesis (DEC) model of range evolution, consisting of 15 component areas, using the program Lagrange (Ree and Smith 2008). The areas, shown in Fig. 1, were taken from Mansion et al. (2008), who defined them on the basis of continental plates or microplates identified in tectonic reconstructions of the Mediterranean basin (Stampfli et al. 1991; Rögl 1999; Krijgsman 2002; Rosenbaum et al. 2002a; Meulenkamp and Sissingh 2003). Species ranges were defined by presence-absence coding (online Appendix 4). The DEC model describes ancestor-descendant transitions between geographic ranges by processes of dispersal (range expansion), local extinction (range contraction), and cladogenesis (range subdivision/inheritance), and defines the likelihood of species-range data arrayed at the tips of a phylogenetic tree as a function of rates of dispersal and local extinction (see Ree and Smith 2008). Fifteen areas yield a set of $2^{15} = 32{,}768$ theoretically possible geographic ranges (area subsets), many of which were excluded from consideration based on the biological implausibility of their spatial configurations (e.g., wide disjuncts). We defined biologically plausible ranges as those that were "contiguous": that is, represented connected nodes in an area-adjacency graph. We also discarded ranges larger than two areas in size that were not subsets of observed species ranges (e.g., *R. montana*, with five areas, and *R. chalepensis*, with 10 areas). The motivation for this step was to further reduce the dimensions of the model's transition matrix, thus increasing its computational feasibility. The final number of ranges in the analysis was 412.

Following principles described in Ree and Smith (2008), we constructed temporal constraints on rates of dispersal between areas based on palaeo-geographic reconstructions of area position through time (Carracedo et al. 1998; Anguita and Hernán 2000; Dercourt et al. 2000; Scotese 2001; Rosenbaum et al. 2002a; Meulenkamp and Sissingh 2003). These constraints were implemented as a series of seven time slices. For each, we constructed a matrix of scaling factors (between zero and one) for the dispersal rate between areas according to their geographic position, interpreting greater distances and/or the extent of geographic barriers (e.g., sea straits, mountain chains) as being inversely proportional to the expected rate. All time slices together spanned the past 32.5 MY, with the most recent slice being 2.5 MY in duration, and the rest 5 MY each (Fig. 2 and online Appendix 5).

Analyses using the temporally-constrained DEC model were performed on: i) the maximum-clade-credibility tree inferred with BEAST; ii) the chronogram inferred with MULTIDIVTIME using the 50% majority rule consensus tree, obtained with MRBAYES, as input tree; and iii) the chronogram inferred with MULTIDIVTIME using the MAP tree, obtained with MRBAYES, as input tree. For each tree, "global" rates of dispersal and local extinction were first estimated without conditioning on any particular ancestral areas at internal nodes. Then, these rates were used to find the maximum-likelihood values of ancestral range, and subdivision-inheritance scenarios at each internal node on the tree. Each node was considered in isolation, without conditioning on any other ancestral states.

Results

Phylogenetic analyses

The combined cpDNA matrix included a total of 150 sequences, of which 66 were newly generated (GenBank accession numbers are reported in online Appendix 1). The MP analysis yielded 52 most parsimonious trees of 2815 steps, with a CI of 0.716 and a RI of 0.839. The strict consensus of these trees was topologically identical to the 50% majority rule consensus tree found with BI and to the maximum-clade-credibility tree inferred with BEAST (Fig. 2).

The topology of the BEAST tree was generally well supported in terms of both BP (obtained from both MP and ML analyses) and PP (obtained from the MRBAYES analysis) values, with 12 nodes with BP values < 85 and 7 nodes with PP values < 0.95, out of 43 nodes (Fig. 2). As in Salvo et al. (2008), both the C-S endemics – *R. corsica* and *R. lamarmorae* – and the Canarian endemics – *R. pinnata*, *R. oreojasme*, and *R. microcarpa* – formed strongly-supported clades (nodes d and f; BP = 100, PP = 1; Fig. 2). Most importantly, the nodes of interest (nodes a to g, Fig. 2) and the calibration nodes (nodes 1 to 4, Fig. 2) received high

support in terms of both BP and PP values, except for node e (i.e., the split between the Canarian endemics and their sister group) and node 3 (i.e., the node to which the fossil *Ptelea enervosa* was attached).

Molecular dating analyses

The two separate runs for each MULTIDIVTIME analysis converged on similar posterior distributions. The mean nodal ages and credibility intervals inferred for the nodes of interest are shown in Table 1. The sensitivity analyses demonstrated that only the choice of "bigtime" prior affected the resulting age estimates, but with only very small effects on the nodes of interest (online Appendix 3).

The BEAST analyses found a high level of substitution-rate variation across the sampled sequences, as indicated by the marginal posterior probability of the coefficient of variation of rates: mean = 0.986; 95% highest posterior density (HPD) interval = (0.758, 1.228). Additionally, no evidence of rate autocorrelation between neighbouring branches was detected, as indicated by the marginal posterior probability of rate covariance (mean = -0.0151; 95% HPD interval = (-0.177, 0.177)), although this statistic has been found to have relatively low power (Moore and Donoghue 2007; Ho 2009). The nodal heights and 95% HPD intervals inferred with BEAST are graphically shown in Fig. 2; for the nodes of interest these values are reported in Table 1. The sensitivity analyses indicated that the choice of different parametric distributions for the calibration nodes did not have an effect on the resulting mean nodal heights and 95% HPD intervals (online Appendix 3). The specification of a root prior, instead, yielded younger mean nodal heights and narrower 95% HPD intervals, as compared to analyses carried out without specifying such a prior (online Appendix 3). We decided to conduct the final BEAST analysis implementing a root prior because this produced narrower 95% HPD intervals, and because we did not see any reason why available estimates of the age of the root (i.e., Rutaceae) from the fossil record (Knobloch and Mai 1986) and from previous phylogenetic findings (Muellner et al. 2007) should have been excluded. As in the MULTIDIVTIME analysis, the analysis performed omitting the fossil *Ptelea enervosa* yielded similar age estimates to the analysis carried out including this fossil.

Generally, the mean nodal ages estimated with MULTIDIVTIME were older than those inferred with BEAST (black regression line; online Appendix 6). However, when we compared both mean nodal ages (grey regression line; online Appendix 6) and credibility intervals (Table 1) for the nodes within *Ruta* only, the results of both dating methods converged on similar values.

Ancestral range reconstruction analyses

For each node in each of the three chronograms inferred using MULTIDIVTIME and BEAST, Lagrange estimated the likelihoods of all possible ancestral range reconstructions (online Appendix 7). We used the maximum likelihood values for further analysis, acknowledging that estimation error associated with statistical uncertainty leaves the door open to alternative interpretations. Lagrange analyses of the BEAST chronogram produced a higher likelihood score and global dispersal rate than analyses of the MULTIDIVTIME chronograms (online Appendix 7). Across trees, reconstructions were identical for the initial diversification of *Ruta*, the split between *R. corsica* and *R. lamarmorae*, the origin of the Canarian endemics, the split between *R. oreojasme* and the rest, and the split between *R. pinnata* and *R. microcarpa*, but differed for the origin of *Ruta* and the C-S lineage (Table 1). However, in the last two cases, the ancestral areas inferred with the MULTIDIVTIME trees were always a subset of the areas inferred with the BEAST tree; hence biogeographic scenarios were based on the reconstructions obtained with the latter tree. Generally, nodes that are separated by long branches, such as those inferred with MULTIDIVTIME, correspond to a more restricted range, because longer branches allow more time for dispersal events, which can widen the range, to happen (see Smith 2009).

Discussion

The ancestor of Ruta *invaded the Mediterranean region from the North before the onset of the Mediterranean climate*

The molecular dating and ancestral range analyses suggested that *Ruta* diverged from its sister group, comprising the East Asian *Boenninghausenia* and the disjunct Afro-American *Thamnosma*, in the middle Eocene, in an area comprising Eurasia and western North America (node a; Fig. 3, I, A; Table 1). At that time, the Bering Land Bridge (Marincovich and Gladenkov 2001) might have allowed the ancestor of *Ruta* to attain such a distribution. In the middle Eocene, southern Europe was still an archipelago and the Atlantic and Indian Oceans were connected via the Tethys Sea (Rögl 1999; Fig. 3, A). When *Ruta* started to diversify, in the early Miocene, the palaeo-geographic setting of the developing Mediterranean basin had dramatically changed (node b; Fig. 3, I, B; Table 1). Two key tectonic events occurred between the middle Eocene and early Miocene, which might have facilitated the major southward range expansion of the genus from Eurasia to the margins of the proto-Mediterranean basin, as suggested by the ancestral range reconstruction analyses (Table 1). The progressive accretion of the microplates located between the Paratethys and Tethys Seas (Fig. 3, A) caused the formation of a more-or-less continuous landmass, extending from the

proto-Iberian peninsula to Asia Minor, in the late Oligocene-early Miocene, roughly between 25 and 20 MA (Fig. 3, B; Rögl 1999; Meulenkamp and Sissingh 2003). This new geological configuration allowed biotic exchange between the eastern and western proto-Mediterranean basin, resulting in the circum-Mediterranean distribution of many terrestrial organisms (e.g., butterflies, crane flies, scorpions, frogs, newts, beetles, Asteraceae, *Cyclamen*; Steininger et al. 1985; Oosterbroek and Arntzen 1992; Palmer and Cambefort 2000; Sanmartín 2003; Oberprieler 2005; Mansion et al. 2008; Micó et al. 2009; Yesson and Culham 2009). This landmass might have facilitated land migration of the ancestor of *Ruta* around the proto-Mediterranean basin. Around 20 MA, the collision of the African and Arabian plates with the Anatolian microplate, resulting in the interruption of the Tethys Sea, caused the formation of a land corridor between Africa and Eurasia across Arabia and Asia Minor (Hallam 1981; Rögl 1999; Krijgsman 2002; Lomolino et al. 2006). This almost-continuous land corridor, which was crucial for the Eurasian-African faunal exchange during the early Miocene (Coryndon and Savage 1973; Steininger et al. 1985), likely facilitated the expansion of the ancestor of *Ruta* from Eurasia, through the Arabian plate, to North Africa (Fig. 3, B).

The temporal envelopes for the origin of *Ruta* (node a) and its subsequent invasion of the Mediterranean basin (node b; Fig. 3, I; Table 1) precede both the trend towards increasing aridification, starting 9-8 MA (Ivanov et al. 2002; Van Dam 2006), and the onset of the Mediterranean climate (Suc 1984; Thompson 2005). Although the onset of Mediterranean-type climates has been shown to trigger radiations in some plant groups (e.g., *Pelargonium*, Cape region, Bakker et al. 2005; *Drosera*, southern Australia, Yesson and Culham 2006), no temporal overlap between this climatic event and diversification has been detected in others (e.g., *Anthemis*, Mediterranean region, Lo Presti and Oberprieler 2009; *Protea*, Cape region, Barraclough and Reeves 2005). Perhaps, in the latter instances, the filtering of elements from pre-existing regional species pools and the arrival of others by dispersal from surrounding regions might have prevented the *in situ*, climatically-driven diversification of lineages (Donoghue 2008). In fact, it has been suggested that the filtering of elements from the ancient geofloras that spread across the Northern Hemisphere during the Tertiary (Tertiary geofloras; Wolfe 1975, 1978; Hickey et al. 1983; Tiffney 1985) played a crucial role in the assembly of the Mediterranean floristic diversity (Thompson 2005; Ackerly 2009). For example, Mai (1989) concluded that the elements characteristic of the Mediterranean sclerophyll vegetation originated in the Tertiary geofloras. In particular, taxa such as *Quercus* and *Pistacia*, commonly associated with *Ruta* (Bonet 1992), were derived from the mixed mesophytic forest, a deciduous forest that formed in the Northern Hemisphere during the Oligocene due to the mixing of two Tertiary geofloras, the boreotropical forest and the Arcto-Tertiary geoflora

(Tiffney 1985; Mai 1989). Interestingly, the inferred ancestral area for the origin of *Ruta* (Fig. 3, I, A) overlaps with the proposed range of the mixed mesophytic forest both in space and time (Kvaček et al. 2006; Kürschner and Kvaček 2009). Unfortunately, the poor fossil record of *Ruta* (Gregor 1989) fails to provide any empirical evidence for or against the inferred northern origin of the genus.

In summary, the results of our integrated molecular dating and ancestral range reconstruction analyses indicate that *Ruta* originated *ex situ* (i.e., outside of the Mediterranean region), in a large area comprising Eurasia and western North America, well before the onset of the Mediterranean climate (Fig. 3, I, A). At that time the Northern Hemisphere was covered predominantly by the boreotropical forest (Axelrod 1975; Mai 1989; Thompson 2005). It then expanded its range southward, invaded several landmasses situated around the forming Mediterranean basin, (Fig. 3, I, B), and diversified *in situ* as the climate changed from subtropical to Mediterranean. Such diversification was probably driven by the geologic complexity of the region, characterized by the appearance and disappearance of barriers to dispersal, as found for other Mediterranean groups (Palmer and Cambefort 2000; Sanmartín 2003; Oberprieler 2005; Thompson 2005; Mansion et al. 2008).

The origin of the Corso-Sardinian lineage is better explained by the separation between the C-S and Apulian microplates in the Miocene than by the fragmentation of the Hercynian massif in the Oligocene

Island endemics have been traditionally divided into two groups: palaeo-endemics and neo-endemics (Favarger and Contandriopoulos 1961; Stebbins and Major 1965; Bramwell 1972; Mansion et al. 2009). The former are ancient lineages, often relict elements of once-widespread groups, which are geographically and taxonomically isolated from their closest extant relatives and show little geographic variation; the latter are more recently evolved and exhibit geographical and taxonomic proximity to their closest extant relatives (Cardona and Contandriopoulos 1979; Thompson 2005). Due to their narrow distribution – being present in only a few, isolated mountains of Corsica and Sardinia – and their morphological distinctness, as compared to other congeneric species (Bacchetta et al. 2006), *R. corsica* and *R. lamarmorae* have long been considered as relictual, palaeo-endemic species (Cardona and Contandriopoulos 1979; Arrigoni 1983; Thompson 2005), their origin being linked to the Oligocene separation of the C-S microplate from the proto-Iberian peninsula (Gamisans 1975).

The molecular dating and ancestral range reconstruction analyses indicated that the ancestor of the two C-S endemics and their widespread sister clade occurred in a broad area, ranging from Eurasia to North Africa, during the middle Miocene (node c; Fig. 3, II, C; Table

1). At that time, the C-S microplate was connected to the Apulian microplate (Fig. 3, C; Cherchi and Montadert 1982; Deino et al. 2001; Rosenbaum et al. 2002a; Speranza et al. 2002; Rosenbaum and Lister 2004), suggesting an invasion of the former via land migration through the latter. The subsequent separation between these two landmasses might have driven the allopatric divergence of the ancestor of *R. corsica* and *R. lamarmorae* from the mainland relatives. A visual inspection of the chronogram (Fig. 3, II) indicates that the time window associated with the connection between the C-S and Apulian microplates is much larger than the time windows associated with other events. This is the result of the uncertainty surrounding this tectonic event (Robertson and Grasso 1995). In particular, the formation of the Tyrrhenian Sea, which terminated the above-mentioned connection, is one of the most complicated aspects of the geologic history of the Mediterranean basin, since it is the result of the interaction of three mountain chains – the Alps, the Apennines, and the Maghrebides (Selli 1985). Nevertheless, the envelope of uncertainty surrounding the above-mentioned connection overlaps with the credibility interval for the age of the split between the ancestor of the C-S endemics and their sister clade (Fig. 3, II), strongly suggesting that both processes of land migration through the temporary C-S block/Apulia land corridor and allopatric speciation driven by the disruption of such corridor played a role in the origin of the C-S endemic lineage. These results emphasize the role of land connections between the C-S microplate and adjacent areas during the migration of the C-S block towards its current position in the assembly of the C-S flora, as first suggested by Braun-Blanquet (1926).

Mansion et al. (2009) proposed an explicit spatio/temporal approach to discriminate between island palaeo- and neo-endemics. If a speciation event post-dated the formation of either an oceanic or continental fragment island, resulting in the *in situ* divergence of the island endemic from its closest relative, then the island endemic was classified as "neo-endemic". If the speciation event is driven by or precedes island formation, the island endemic was classified as "palaeo-endemic", i.e., either a surviving element of the ancestral continental biota on the island or an immigrant that might have originated *ex situ* and colonized the island after its formation via LDD. With respect to continental fragment islands, it is important to pinpoint the geologic event that led to the separation from the continent, which can be used as temporary boundary for the classification of the islands' endemic species as either palaeo- or neo-endemic. While the connection of Corsica and Sardinia with the proto-Iberian peninsula has been known for a relatively long time (Alvarez et al. 1974; Boccaletti and Guazzone 1974), the subsequent geologic events involving the interaction of the two islands with the Apulian microplate have been elucidated only recently and are more controversial (Rosenbaum and Lister 2004; Rosenbaum et al. 2008). Thus biogeographers

have traditionally focused on the fragmentation of the Hercynian massif (30-28 MA) to classify the C-S endemic flora into palaeo- or neo-endemic species (e.g., Gamisans 1975; Cardona and Contandriopoulos 1979), rather than on subsequent tectonic events. However, since the credibility interval for the divergence between the C-S endemic lineage and its closest relatives overlaps with the temporal window of the separation between the C-S and Apulian microplates (Fig. 3, II), the continental fragment island endemics can still be considered palaeo-endemics, but in relation to this later geologic event, rather than the intial splitting of the C-S microplate from the proto-Iberian peninsula.

The split between the Corsican and Sardinian endemics is better explained by events associated with the Messinian Salinity Crisis than by the formation of the Strait of Bonifacio

Salvo et al. (2008) conjectured that the divergence between *R. corsica* and *R. lamarmorae* might have been caused by vicariance driven by the geologic formation of the Strait of Bonifacio in the middle Miocene (15-9 MA; Alvarez et al. 1974; Bonin et al. 1979; Orsini et al. 1980; Cherchi and Montadert 1982). However, subsequent climatic events might have also caused such divergence, for example, the end of the MSC (Rögl and Steininger 1983; Gover et al. 2009) and the end of the LGM (Lambeck et al. 2004; Lambeck and Purcell 2005), which marked the re-flooding of the Strait of Bonifacio.

The ancestral range reconstruction analyses inferred a potential area of distribution for the ancestor of *R. corsica* and *R. lamarmorae* comprising Corsica and Sardinia (node d; Fig. 3, III, D; Table 1), thus congruent with a scenario of vicariant speciation. However, the molecular dating analyses estimated an age of 4-3 MA for the split between the two species (Fig. 3, III; Table 1). This estimate and its credibility interval post-date the formation of the Strait of Bonifacio, indicating that the divergence between the two C-S endemics was not associated with this tectonic event. Instead, the credibility interval for the split between *R. corsica* and *R. lamarmorae* overlaps with the end of the MSC (5.33 MA), when the flooding of the Mediterranean basin caused the renewed separation between the two islands (Fig. 3, III).

The MSC has been repeatedly invoked to explain current distributional patterns of different Mediterranean taxa (Bocquet et al. 1978; Palmer and Cambefort 2000; Sanmartín 2003; Mansion et al. 2009). The drying out and re-flooding of the Mediterranean basin is a plausible mechanism by which barriers to terrestrial dispersal were removed and created, respectively (Yesson et al. 2009). Altaba (1998), however, cast doubt on the biogeographical significance of this event, since: (i) the palaeo-geography of the Mediterranean basin during the MSC is still unclear (Meijer and Krijgsman 2005), (ii) dispersal across a hot, saline desert

may have been difficult for most organisms, and (iii) most biogeographic studies of the MSC have found a link between differentiation in terrestrial taxa and the onset of the MSC, whereas such differentiation should be connected with the re-filling of the basin.

The two C-S endemics are restricted to the main mountainous areas of Corsica and Sardinia, growing above 1000 m. a. s. l. (Fig. 1). Moreover, they are very specialized in their habitat requirements, occurring mainly on siliceous substrates and amid the *Carici-Genistetea lobelii* plant community (Bacchetta et al. 2006). With these facts in mind, an interesting line of future research concerns the reconstruction of ancestral ecological niches, by using the techniques of phyloclimatic modelling (Yesson and Culham 2006), in order to determine which parts of the inferred ancestral range (i.e., Corsica and Sardinia) might have been environmentally suitable.

The colonization of the Canary Islands by the Canarian clade of* Ruta *does not conform to a stepping-stone model and its origin predated the onset of the Mediterranean climate

The phylogenetic analyses showed that the Canarian endemics – *R. pinnata*, *R. oreojasme*, and *R. microcarpa* – form a monophyletic group (node f; Fig. 2), as in the majority of Canarian genera so far investigated (Carine et al. 2004), indicating that they were the product of a single colonization event into the islands (Fig. 4). The ancestral range reconstruction analyses inferred North Africa to be the source area of such colonization event (node e; Fig. 4; Table 1), as commonly found in other groups (Carine et al. 2004). Due to the dynamic geologic history of the Canarian archipelago (Whittaker et al. 2008), biogeographic scenarios involving numerous inter-island dispersal events, also to islands that are no longer emerged, are possible. The most parsimonious interpretation of the inferred ancestral areas and present distribution of the Canarian endemics requires a minimum of four dispersal events: one from North Africa (the ancestral area inferred for node e) to Gomera (the ancestral area inferred for node f); one from Gomera to Gran Canaria; one from Gomera to Tenerife (one of the ancestral areas inferred for node g); and one from Tenerife to La Palma (Fig. 4, right; Table 1).

By integrating evidence from the molecular dating results and the geologic history of the Canary Islands, we are able to add a temporal dimension to the hypothesized dispersal events. Since these events are inferred along internodes, the temporal uncertainty associated with the nodes bracketing the internodes in question has to be taken into account. For example, we hypothesize a temporal envelope for dispersal event 1, which was inferred along the internode connecting nodes e and f, spanning the lower bound of the credibility interval (CI) of node e (27.3 MA) and the upper bound of the CI of node f (2.6 MA; Fig. 4, left; Table 1). We proceed in a similar fashion for the inferred dispersal events 2, 3, and 4 (Fig. 4).

Overlap between island emergence (Carracedo et al. 1998; Anguita and Hernán 2000) and temporal envelopes for each dispersal event strongly suggests that the most likely time windows of island colonization were 12-2.6 MA for Gomera, 14.9-0.3 MA for Gran Canaria, 11.6-0.5 MA for Tenerife, and 2-0 MA for La Palma (Fig. 4, left). Since Gomera was the source area for the colonization of Gran Canaria, dispersal event 2 could not have predated dispersal event 1.

Three main modes of species diversification in the Canary Islands have been identified: i) stepping-stone, with a single colonization event from the mainland followed by colonization events proceeding from older to younger islands; ii) multiple independent colonization events from the mainland, followed by within-island speciation; and iii) inter-island colonization between similar ecological habitats (Sanmartín et al. 2008). For the Canarian endemics of *Ruta*, only one colonization event from the mainland was inferred and, hence, mode of diversification ii can be ruled out.

Since Gran Canaria is older, and closer to the mainland, than Gomera (Carracedo et al. 1998; Anguita and Hernán 2000), and since the sequence of splitting events within the Canarian clade is congruent with the order of island formation, a stepping-stone mode of island colonization seemed plausible. However, Gomera, and not Gran Canaria, was inferred as the ancestral range for node f with both BEAST and MULTIDIVTIME chronograms, suggesting that Gomera was the first island to be colonized by *Ruta*, and casting doubt on a stepping-stone biogeographic scenario (Fig. 4). This counterintuitive result might depend on the uncertainty surrounding the Lagrange reconstructions (online Appendix 7; see also Ree and Smith 2008) and the complexity of the underlying temporally-constrained DEC model (e.g., Clark et al. 2008). Therefore, our proposed pathways for the colonization of the Canary Islands should be viewed as a starting hypothesis that can be tested in future studies based on expanded infra-specific and inter-island taxon sampling.

R. pinnata, *R. oreojasme*, and *R. microcarpa* are confined to the same vegetational zone, thermophilous scrubland (Bramwell and Bramwell 2001), which is a recent ecosystem, believed to have originated concomitantly with the onset of the Mediterranean climate (~3 MA; Fernández-Palacios et al. 2008). For this reason, these species were assumed to have originated together with the onset of such climate. However, the inferred temporal windows for the intial invasion of the Canary archipelago and subsequent island colonizations (i.e., except for La Palma) precede the onset of the Mediterranean climate (Fig. 4, bottom left), and overlap with a time when the Canary Islands were mainly covered by laurel and pine forests (Fernández-Palacios et al. 2008). This suggests that i) when *Ruta* colonized the Canary Islands, the islands' vegetation was very different than at present; ii) in order to persist under

the novel Mediterranean climate/vegetation, *Ruta* likely changed its ecological requirements; iii) the filtering of the Canarian endemics into thermophilous scrublands after the onset of the Mediterranean climate occurred in parallel in different islands; and iv) inter-island colonization between similar ecological habitats cannot be readily endorsed, because it does not take into account the different time windows inferred for the colonization of the different islands. Again, the application of ecological niche modelling approaches within a phylogenetic framework might help to elucidate some of these issues.

Conclusion

Our study demonstrates that the integration of different sources of information from phylogenetics, molecular dating, ancestral range reconstruction, and geologic/palaeo-climatic models is indispensable for explaining biogeographic patterns. Additionally, the clear formulation of a hypothesis-based framework at the onset of the research helps to avoid the construction of *a posteriori* biogeographic scenarios. With respect to island biogeography theory (MacArthur and Wilson 1967; Whittaker and Fernandez-Palacios 2007; Whittaker et al. 2008), our study stresses the importance of temporary land connections in the biotic assembly of continental fragment islands and of determining discrete time windows of island colonization in order to better understand distributional patterns in oceanic islands (Carine 2005; Kim et al. 2008).

The palaeo-geographic and palaeo-climatic settings in which biodiversity evolved should be carefully incorporated in biogeographic studies, as implemented in recently developed approaches to ancestral range reconstruction (i.e., Lagrange; Ree and Smith 2008). Inferred areas, however, can only indicate the maximum possible extent of the distribution of an ancestor. To achieve more realistic estimates of ancestral distributions, methods for the projection of ancestral ecological niches into palaeo-climatic and palaeo-geographic configurations need to be further developed (Yesson and Culham 2006; Evans et al. 2009). Within *Ruta*, the application of niche modelling tools will be fundamental to achieve a more complete understanding of the relative roles of geologic versus climatic factors in speciation and of niche conservatism versus niche evolution in shaping distributional patterns in the Mediterranean region (Donoghue 2008; Ackerly 2009).

Funding

This work was supported by the Swiss National Science Foundation grant no. 3100A0-105495 to E. C. and the Institute of Systematic Botany of the University of Zürich. Additional financial support from the Georges-und-Antoine-Claraz-Schenkung, the

Kommission für Reisestipendien der Schweizerischen Akademie der Naturwissenschaften (SANW), and the Equal Opportunity Office of the University of Zurich are gratefully acknowledged.

Acknowledgements

The authors would like to thank: B. Keller, J. de Vos, S. Manafzadeh, L. Granato – help and comments on an early draft; G. Bacchetta, A. Bertini, J. M. Fernández-Palacios, W. Krijgsman, A. Antonelli, S. Magallón – stimulating discussions; G. Mansion – help in the field; M. Pirie – help with r8s and BEAST; J. Fuertes-Aguilar, A. Santos-Guerra – material from the Canary Islands; L. Hugot – collecting permits in Corsica; B. Seitz, E. Benetti – help in the library; Jack Sullivan, Dan Faith, Isabel Sanmartín, one anonymous reviewer, and Deborah Ciszek – constructive criticism on the manuscript.

References

Ackerly, D. 2009. Some comments on the age, origin and evolution of the California and Mediterranean floras. J. Biogeogr. 36:1221–1233.

Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Contrology 19:716–723.

Altaba, C. R. 1998. Testing vicariance: melanopsid snails and Neogene tectonics in the western Mediterranean. J. Biogeogr. 25:541–551.

Alvarez, W. 1972. Rotation of the Corsica-Sardinia microplate. Nature Phys. Sci. 235:103–105.

Alvarez, W. 1976. A former continuation of the Alps. Geol. Soc. Am. Bull. 87:891–896.

Alvarez, W., T. Cocozza, and F. C. Wezel. 1974. Fragmentation of Alpine orogenic belt by microplate dispersal. Nature 248:309–314.

Anderson, C. L., K. Bremer, and E. M. Friis. 2005. Dating phylogenetically basal eudicots using *rbcL* sequences and multiple fossil reference points. Am. J. Bot. 92:1737–1748.

Anguita, F., and F. Hernán. 2000. The Canary Islands origin: a unifying model. J. Volcanol. Geoth. Res. 103:1–26.

Arrigoni, P. V. 1979. Le piante endemiche della Sardegna. Boll. Soc. Sarda Sci. Nat. 18:223–295.

Arrigoni, P. V. 1983. Aspetti corologici della flora sarda. Lav. Soc. Ital. Biogeogr. 8:83–109.

Axelrod, D. I. 1975. Evolution and biogeography of Madrean-Tethyan sclerophyll vegetation. Ann. Mo. Bot. Gard. 62:280–334.

Bacchetta, G., S. Brullo, and G. Giusso del Galdo. 2006. *Ruta lamarmorae* (Rutaceae), a new species from Sardinia. Edinb. J. Bot. 63:153–160.

Bakker, F. T., A. Culham, E. M. Marais, and M. & Gibby. 2005. Nested radiation in Cape *Pelargonium*. Pages 75–100 *in* Plant species-level systematics: new perspectives on pattern and process (F. T. Bakker, L. W. Chatrou, B. Gravendeel, and P. Pelser, eds.). A. R. G. Gantner Verlag, Ruggell, Liechtenstein.

Barbey, W. 1885. Florae sardoae compendium. Catalogue raisonné des végétaux observés dans l'Île de Sardaigne. Bridel, Lausanne.

Barraclough, T. G., and G. Reeves. 2005. The causes of speciation in flowering plant lineages: species-level DNA trees in the African genus *Protea*. Pages 31–46 *in* Plant species-level systematics: new perspectives on pattern and process (F. T. Bakker, L.

W. Chatrou, B. Gravendeel, and P. Pelser, eds.). A. R. G. Gantner Verlag, Ruggell, Liechtenstein.

Baum, D. A., R. L. Small, and J. F. Wendel. 1998. Biogeography and floral evolution of baobabs (*Adansonia*, Bombacaceae) as inferred from multiple data sets. Syst. Biol. 47:181–207.

Boccaletti, M., and G. Guazzone. 1974. Remnant arcs and marginal basins in the Cainozoic development of the Mediterranean. Nature 252:18–21.

Bocquet, G., B. Widler, and H. Kiefer. 1978. The Messinian Model - A new outlook for the floristics and systematics of the Mediterranean area. Candollea 33:269–287.

Bonet, A. 1992. Biología reproductiva de *Ruta angustifolia* Pers. en la cordillera litoral catalana. Fol. Bot. Misc. 8:113–124.

Bonin, B., P. Chotin, A. Girte, and J. Orsini. 1979. Étude du bloc corso-sarde sur documents satellites: le problème des mouvements différentiels entre les deux iles. Rev. Geog. Phys. Geol. Dynam. 21:147–154.

Bramwell, D. 1972. Endemism in the flora of the Canary Islands. Pages 141–159 *in* Taxonomy, phytogeography and evolution (D. H. Valentine, ed.) Academic Press, London.

Bramwell, D., and Z. I. Bramwell. 2001. Wild flowers of the Canary Islands, 2nd edition. Editorial Rueda, Madrid.

Braun-Blanquet, J. 1926. Les phanérogames, histoire du peuplement de la Corse. Bull. Soc. Hist. Nat. Corse 45:1–17.

Briquet, J. 1910. Prodrome de la flore Corse. Georg & Cie, Genève, Bâle et Lyon.

Brown, J. M., and A. R. Lemmon. 2007. The importance of data partitioning and the utility of Bayes factors in Bayesian phylogenetics. Syst. Biol. 56:643–655.

Caccone, A., M. C. Milinkovitch, V. Sbordoni, and J. R. Powell. 1994. Molecular biogeography: using the Corsica-Sardinia microplate disjunction to calibrate mitochondrial rDNA evolutionary rates in mountain newts (*Euproctus*). J. Evol. Biol. 7:227–245.

Caccone, A., and V. Sbordoni. 2001. Molecular biogeography of cave life: a study using mitochondrial DNA from Bathysciine beetles. Evolution 55:122–130.

Call, V. B., and D. L. Dilcher. 1995. Fossil *Ptelea* samaras (Rutaceae) in North America. Am. J. Bot. 82:1069–1073.

Cardona, M. A., and J. Contandriopoulos. 1979. Endemism and evolution in the islands of the western Mediterranean. Pages 133–169 *in* Plants and islands (D. Bramwell, ed.) Academic Press, London.

Carine, M. A. 2005. Spatio-temporal relationships of the Macaronesian endemic flora: a relictual series or window of opportunity? Taxon 54:895–903.

Carine, M. A., S. J. Russell, A. Santos-Guerra, and J. Francisco-Ortega. 2004. Relationships of the Macaronesian and Mediterranean floras: molecular evidence for multiple colonizations into Macaronesia and back-colonization of the continent in *Convolvulus* (Convolvulaceae). Am. J. Bot. 91:1070–1085.

Carracedo, J. C., S. Day, H. Guillou, E. R. Badiola, J. A. Canas, and F. J. P. Torrado. 1998. Hotspot volcanism close to a passive continental margin: the Canary Islands. Geol. Mag. 135:591–604.

Cavazza, W., F. M. Roure, W. Spakman, G. Stampfli, and P. A. Ziegler. 2004. The TRANSMED Atlas - The Mediterranean region from crust to mantle. Springer, Berlin, Heidelberg.

Chase, M. W., C. M. Morton, and J. A. Kallunki. 1999. Phylogenetic relationships of Rutaceae: a cladistic analysis of the subfamilies using evidence from *rbc*L and *atp*B sequence variation. Am. J. Bot. 86:1191–1199.

Cherchi, A., and L. Montadert. 1982. Oligo-Miocene rift of Sardinia and the early history of the western Mediterranean basin. Nature 298:736–739.

Clark, J. R., R. H. Ree, M. E. Alfaro, M. G. King, W. L. Wagner, and E. H. Roalson. 2008. A comparative study in ancestral range reconstruction methods: retracing the uncertain histories of insular lineages. Syst. Biol. 57:693–707.

Contandriopoulos, J. 1962. Recherches sur la flore endémique de la Corse et sur ses origines. Faculté des Sciences, Montpellier.

Conti, E., F. Rutschmann, T. Eriksson, K. J. Sytsma, and D. A. Baum. 2004. Response to Robert G. Moyle "Calibration of molecular clocks and the biogeographic history of Crypterionaceae". Evolution 58:1874–76.

Coryndon, S. C., and R. J. G. Savage. 1973. The origin and affinities of African mammal faunas. Spec. Pap. Paleontol. 12:121–135.

Cowie, R. H., and B. S. Holland. 2006. Dispersal is fundamental to biogeography and the evolution of biodiversity in oceanic islands. J. Biogeogr. 33:193–198.

Crisci, J. V., L. Katinas, and P. Posadas. 2003. Historical Biogeography: an introduction. Harvard University Press, Cambridge, MA.

Da Silva, M. F., O. R. Gottlieb, and F. Ehrendorfer. 1988. Chemosystematics of the Rutaceae: suggestions for a more natural taxonomy and evolutionary interpretation of the family. Plant Syst. Evol. 161:97–134.

Deino, A., J. Gattacceca, R. Rizzo, and A. Montanari. 2001. $^{40}Ar/^{39}Ar$ dating and paleomagnetism of the Miocene volcanic succession of Monte Furru (western Sardinia): implications for the rotation history of the Corsica-Sardinia microplate. Geophys. Res. Lett. 28:3373–3376.

Dercourt, J., M. Gaetani, B. Vrielynck, E. Barrier, B. Biju-Duval, M. F. Brunet, J. P. Cadet, S. Crasquin, and M. e. Sandulescu. 2000. Atlas Peri-Tethys, Palaeogeographical maps. CCGM/CGMW, Paris.

Dercourt, J., L. P. Zonenshain, L. E. Ricou, V. G. Kazmin, X. Le Pichon, A. L. Knipper, C. Grandjacquet, I. M. Sbortshikov, J. Geyssant, J. Lepvrier, D. H. Pechersky, J. Boulin, J. C. Sibuet, L. A. Savostin, O. Sorokhtin, M. Westphal, M. L. Bazhenov, J. P. Lauer, and B. Biju-Duval. 1986. Geological evolution of the Tethys belt from the Atlantic to the Pamirs since the Lias. Tectonophysics 123:241–315.

Dewey, J. F., M. L. Helman, E. Turco, D. H. W. Hutton, and S. D. Knott. 1989. Kinematics of the western Mediterranean. Pages 265–283 *in* Alpine Tectonics (M. P. Coward, D. Dietrich, and R. G. Park, eds.). Geological Society, London, Special Publications, v. 45.

Donoghue, M. J. 2008. A phylogenetic perspective on the distribution of plant diversity. PNAS 105:11549–11555.

Donoghue, M. J., and B. R. Moore. 2003. Toward an integrative historical biogeography. Integr. Comp. Biol. 43:261–270.

Drummond, A. J., S. Y. W. Ho, M. J. Phillips, and A. Rambaut. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4:e88.

Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7:214.

Emerson, B. C. 2002. Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. Mol. Ecol. 11:951–966.

Engler, A. 1879. Versuch einer Entwicklungsgeschichte der Pflanzenwelt, insbesondere der Florengebiete seit der Tertiärperiode 1. Engelmann, Leipzig.

Engler, A. 1896. Rutaceae. Pages 95–201 *in* Nat. Pflanzenfam III, Vol. 4 (A. Engler, and K. Prantl, eds.). Engelmann, Leibzig.

Engler, A. 1936. Rutaceae. Pages 187–359 *in* Die natürlichen Pflanzenfamilien, 2 Aufl 19a (A. Engler, and K. Prantl, eds.). Engelmann, Leibzig.

Evans, M. E. K., S. A. Smith, R. S. Flynn, and M. J. Donoghue. 2009. Climate, niche evolution, and diversification of the "Bird-Cage" evening primroses (Oenothera, sections Anogra and Kleinia). Am. Nat. 173:225–240.

Favarger, C., and J. Contandriopoulos. 1961. Essai sur l'endémisme. Bull. Soc. Bot. Suisse 77:383–408.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the Bootstrap. Evolution 39:783–791.

Fernández-Palacios, J. M., R. Otto, J. D. Delgado, J. R. Arévalo, A. Naranjo, F. González Artiles, C. Morici, and R. Barone. 2008. Los bosques termófilos de Canarias. Proyecto LIFE04/NAT/ES/000064. Cabildo Insular de Tenerife. Santa Cruz de Tenerife.

Fochetti, R., V. Ketmaier, M. Oliverio, J. M. Tierno de Figueroa, and E. Sezzi. 2004. Biochemical systematics and biogeography of the Mediterranean genus *Tyrrhenoleuctra* (Plecoptera, Insecta). Insect Syst. Evol. 35:299–306.

Gadek, P. A., E. S. Fernando, C. J. Quinn, S. B. Hoot, T. Terrazas, M. C. Sheahan, and M. W. Chase. 1996. Sapindales: molecular delimitation and infraordinal groups. Am. J. Bot. 83:802–811.

Gamisans, J. 1975. La végétation des montagnes corses, Aix-Marseille.

Gamisans, J., and D. Jeanmonod. 1993. Catalogue des plantes vasculaires de la Corse (éd 2)*in* Compléments au Prodrome de la Flore Corse (D. Jeanmonod, and H. M. Burdet, eds.). Editions des Conservatoire et Jardin botanique de la Ville de Genève, Genève.

Garfunkel, Z. 2004. Origin of the eastern Mediterranean basin: a reevaluation. Tectonophysics 391:11–34.

Gover, R., P. Meijer, and W. Krijgsman. 2009. Regional isostatic response to Messinian Salinity Crisis events. Tectonophysics 463:109–129.

Gradstein, F. M., J. G. Ogg, A. G. Smith, and e. al. 2004. A Geologic Time Scale 2004. Cambridge University Press, Cambridge.

Gregor, H.-J. 1979. Systematics, biostratigraphy and paleoecology of the genus *Toddalia* Jussieu (Rutaceae) in the European Tertiary. Rev. Palaeobot. Palynol. 28:311–363.

Gregor, H.-J. 1989. Aspects of the fossil record and phylogeny of the family Rutaceae (*Zanthoxyleae*, *Toddalioideae*). Plant Syst. Evol. 162:251–265.

Groppo, M., J. R. Pirani, M. L. F. Salatino, S. R. Blanco, and J. A. Kallunki. 2008. Phylogeny of Rutaceae based on two noncoding regions from cpDNA. Am. J. Bot. 95:985–1005.

Hallam, A. 1981. Relative importance of plate movements, eustasy, and climate in controlling major biogeographical changes since the Early Mesozoic*in* Vicariance biogeography: A critique (G. J. Nelson, and D. E. Rosen, eds.). Columbia University Press, New York.

Heads, M. 2005. Dating nodes on molecular phylogenies: a critique of molecular biogeography. Cladistics 21:62–78.

Hickey, L. J., R. M. West, M. R. Dawson, and D. K. Choi. 1983. Arctic terrestrial biota: paleomagnetic evidence of age disparity with mid-northern latitudes during the Late Cretaceous and early Tertiary. Science 221:1153–1156.

Ho, S. Y. W. 2009. An examination of phylogenetic models of substitution rate variation among lineages. Biol. Lett. 5:421–424.

Ho, S. Y. W., and M. J. Phillips. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst. Biol. 58:367–380.

Hulsenbeck, J. P., and F. Ronquist. 2001. Mr Bayes: Bayesian inference of phylogeny. Bioinformatics 17:754–755.

Hunn, C. A., and P. Upchurch. 2001. The importance of time/space in diagnosing the causality of phylogenetic events: towards a "chronobiogeographical" paradigm? Syst. Biol. 50:391–407.

Ivanov, D., A. R. Ashraf, V. Mosbrugger, and E. Palamarev. 2002. Palynological evidence for Miocene climate change in the Forecarpathian Basin (Central Paratethys, NW Bulgaria). Palaeogeogr. Palaeoclimatol. Palaeoecol. 178:19–37.

Jeanmonod, D., and J. Gamisans. 1987. Introduction*in* Compléments au Prodrome de la Flore Corse (D. Jeanmonod, and H. M. Burdet, eds.). Editions des Conservatoire et Jardin

botanique de la Ville de Genève, Genève.
Juan, C., B. C. Emerson, P. Oromí, and G. M. Hewitt. 2000. Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. Trends Ecol. Evol. 15:104–109.
Ketmaier, V., R. Argano, and A. Caccone. 2003. Phylogeography and molecular rates of subterranean aquatic Stenasellid Isopods with a peri-Tyrrhenian distribution. Mol. Ecol. 12:547–555.
Ketmaier, V., F. Giusti, and A. Caccone. 2006. Molecular phylogeny and historical biogeography of the land snail genus Solatopupa (Pulmonata) in the peri-Tyrrhenian area. Mol. Phylogenet. Evol. 39:439–451.
Kim, S.-C., M. R. McGowen, P. Lubinsky, J. C. Barber, M. E. Mort, and A. Santos-Guerra. 2008. Timing and tempo of early and successive adaptive radiations in Macaronesia. PLoS One 3:e2139. doi:10.1371/journal.pone.0002139.
Kishino, H., J. L. Thorne, and W. J. Bruno. 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. Mol. Biol. Evol. 18:352–361.
Knobloch, E., and D. H. Mai. 1986. Monographie der Früchte und Samen in der Kreide von Mitteleuropa. Rozpr. Ustr. Ust. Geol. 47:1–219.
Krijgsman, W. 2002. The Mediterranean: *mare nostrum* of Earth sciences. Earth Planet. Sc. Lett. 205:1–12.
Kürschner, W. M., and Z. Kvacek. 2009. Oligocene-Miocene CO_2 fluctuations, climatic and palaeofloristic trends inferred from fossil plant assemblages in central Europe. Bull. Geosci. 84:189–202.
Kvacek, Z., M. Kovác, J. Kovar-Eder, N. Doláková, H. Jechorek, V. Parashiv, M. Kovácová, and L. Sliva. 2006. Miocene evolution of landscape and vegetation in the Central Paratethys. Geol. Carpath. 57:295–310.
Lambeck, K., F. Antonioli, A. Purcell, and S. Silenzi. 2004. Sea-level change along the Italian coast for the past 10,000 yr. Quaternary Sci. Rev. 23:1567–1598.
Lambeck, K., and A. Purcell. 2005. Sea-level change in the Mediterranean Sea since the LGM: model predictions for tectonically stable areas. Quaternary Sci. Rev. 24:1969–1988.
Lieberman, B. S. 2003. Unifying theory and methodology in biogeography. Evol. Biol. 33:1–25.
Lo Presti, R. M., and C. Oberprieler. 2009. Evolutionary history, biogeography and eco-climatological differentiation of the genus *Anthemis* L. (Compositae, Anthemideae) in the circum-Mediterranean area. J. Biogeogr. 36:1313–1332.
Lomolino, M. V., B. R. Riddle, and J. H. Brown. 2006. Biogeography; 3rd ed, 3rd edition. Sinauer Associates, Sunderland, MA.
MacArthur, R. H., and E. O. Wilson. 1967. The theory of island biogeography. Princeton University Press, Princeton, NJ.
Maddison, P. G., and D. R. Maddison. 2000. MacClade4: analysis of phylogeny and character evolution. Sinauer, Sunderland, MA.
Magallón, S. A. 2004. Dating lineages: molecular and palaeontological approaches to the temporal framework of clades. Int. J. Plant Sci. 165:S7–S21.
Mai, D. H. 1989. Development and regional differentiation of the European vegetation during the Tertiary. Plant Syst. Evol. 162:79–91.
Mansion, G., G. Rosenbaum, J. Schoenenberger, G. Bacchetta, J. Rosselló, and E. Conti. 2008. Phylogenetic analysis informed by geological history supports multiple, sequential invasions of the Mediterranean basin by the Angiosperm family Araceae. Syst. Biol. 57:269–285.
Mansion, G., F. Selvi, A. Guggisberg, and E. Conti. 2009. Origin of Mediterranean insular endemics in the Boraginales: integrative evidence from molecular dating and ancestral

area reconstruction. J. Biogeogr. 36:1282–1296.
Marincovich Jr., L., and A. Y. Gladenkov. 2001. New evidence for the age of Bering Strait. Quaternary Sci. Rev. 20:329–335.
Mariotti, M. 1990. Floristic connections between the Sardo-Corsican dominion and the Ligurian area. Atti Conv. Lincei 85:429–448.
Médail, F., and P. Quézel. 1997. Hot-spots analysis for conservation of plant biodiversity in the Mediterranean basin. Ann. Mo. Bot. Gard. 84:112–127.
Meijer, P., and W. Krijgsman. 2005. A quantitative analysis of the desiccation and re-filling of the Mediterranean during the Messinian Salinity Crisis. Earth Planet. Sc. Lett. 240:510–520.
Meulenkamp, J. E., and W. Sissingh. 2003. Tertiary palaeogeography and tectonostratigraphic evolution of the Northern and Southern Peri-Tethys platforms and the intermediate domains of the African-Eurasian convergent plate boundary zone. Palaeogeogr. Palaeoclimatol. Palaeoecol. 196:209–228.
Micó, E., I. Sanmartín, and E. Galante. 2009. Mediterranean diversification of the grass-feeding Anisopliina beetles (Scarabaeidae, Rutelinae, Anomalini) as inferred by bootstrap-averaged dispersal-vicariance analysis. J. Biogeogr. 36:546–560.
Moore, B. R., and M. J. Donoghue. 2007. Correlates of diversification in the plant clade Dipsacales: geographic movement and evolutionary innovations. Am. Nat. 170:S28–S55.
Moris, G. G. 1837-59. Flora Sardoa, Taurini, Ex Regio Typographeo.
Muellner, A. N., D. D. Vassiliades, and S. S. Renner. 2007. Placing Biebersteiniaceae, a herbaceous clade of Sapindales, in a temporal and geographic context. Plant Syst. Evol. 266:233–252.
Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature 403:853–858.
Nylander, J. A. A., F. Ronquist, J. P. Huelsenbeck, and J. L. Nieves-Aldrey. 2004. Bayesian phylogenetic analysis of combined data. Syst. Biol. 53:47–67.
Nylander, J. A. A., U. Olsson, P. Alström, and I. Sanmartín. 2008. Accounting for phylogenetic uncertainty in biogeography: a Bayesian approach to Dispersal-Vicariance Analysis of the Thrushes (Aves: *Turdus*). Syst. Biol. 57:257–268.
Oberprieler, C. 2005. Temporal and spatial diversification of circum-Mediterranean Compositae-Anthemideae. Taxon 54:951–966.
Oosterbroek, P., and J. W. Arntzen. 1992. Area-cladograms of circum-Mediterranean taxa in relation to Mediterranean palaeogeography. J. Biogeogr. 19:3–20.
Orsini, J., C. Coulon, and T. Cocozza. 1980. Le dérive cénozoique de la Corse et de la Sardaigne. Géol. Alpine 56:169–202.
Palamarev, D., and K. Usunova. 1970. Morphologisch-anatomischer Nachweis der Gattung Skimmia in der Tertiärflora Bulgariens. C. R. Acad. Bulgare Sci. 23:835–838.
Palmer, M. 1998. Taxonomy, phylogeny, and biogeography of a species-group of West-Mediterranean *Tentyria* (Coleoptera: Tenebrionidae). Ann. Entomol. Soc. Am. 91:260–268.
Palmer, M., and Y. Cambefort. 2000. Evidence for reticulate palaeogeography: beetle diversity linked to connection-disjunction cycles of the Gibraltar Strait. J. Biogeogr. 27:403–416.
Pons, A., and P. Quézel. 1985. The history of the flora and vegetation and past and present human disturbance in the Mediterranean region. Pages 25–43 *in* Plant conservation in the Mediterranean area (C. Gómez-Campo, ed.) Dr. W. Junk Publishers, Dordrecht.
Poon, W.-S., P.-C. Shaw, M. P. Simmons, and P. But. 2007. Congruence of molecular, morphological, and biochemical profiles in Rutaceae: a cladistic analysis of the subfamilies Rutoideae and Toddalioideae. Syst. Bot. 32:837–846.
Posada, D., and K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution.

Bioinformatics 14:817–818.

Quézel, P. 1985. Definition of the Mediterranean region and the origin of its flora. Pages 9–24 *in* Plant conservation in the Mediterranean area. (C. Gómez-Campo, ed.) Dr W. Junk Publishers, Dordrecht.

Ree, R., and I. Sanmartín. 2009. Prospects and challenges for parametric models in historical biogeographic inference. J. Biogeogr. 36:1211–1220.

Ree, R. H., B. R. Moore, C. O. Webb, and M. J. Donoghue. 2005. A Likelihood framework for inferring the evolution of geographic range on phylogenetic trees. Evolution 59:2299–2311.

Ree, R. H., and S. A. Smith. 2008. Maximum-likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 57:4–14.

Reyes-Betancort, J. A., A. Santos-Guerra, I. R. Guma, C. J. Humpries, and M. A. Carine. 2008. Diversity, rarity and the evolution of the Canary Islands endemic flora. An. Jard. Bot. Madr. 65:25–45.

Robertson, A. H. F., and M. Grasso. 1995. Overview of the Late Tertiary: recent tectonic and palaeo-environmental development of the Mediterranean region. Terra Nova 7:114–127.

Rögl, F. 1999. Mediterranean and Paratethys. Facts and hypotheses of an Oligocene to Miocene paleogeography (short overview). Geol. Carpath. 50:339–349.

Rögl, F., and F. F. Steininger. 1983. Vom Zerfall der Tethys zu Mediterran und Paratethys. Die neogene Paläogeographie und Palinspastik des zirkum-mediterranen Raumes. Ann. Nathist. Mus. Wien Ser. B Bot. Zool. 85/A:135–163.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Rosenbaum, G., and G. S. Lister. 2004. Neogene and Quaternary rollback evolution of the Tyrrhenian sea, the Apennines and the Sicilian Maghrebides. Tectonics 23:TC1013, doi:10.1029/2003TC001518.

Rosenbaum, G., G. S. Lister, and C. Duboz. 2002a. Reconstruction of the tectonic evolution of the western Mediterranean since the Oligocene. J. Virt. Ex. 8:107–126.

Rosenbaum, G., G. S. Lister, and C. Duboz. 2002b. Relative motions of Africa, Iberia and Europe during Alpine orogeny. Tectonophysics 359:117–129.

Rosenbaum, G., M.Gasparon, F.P. Lucente, A. Peccerillo, and M.S. Miller. 2008. Kinematics of slab tear faults during subduction segmentation and implications for Italian magmatism. Tectonics 27:TC2008, doi:10.1029/2007TC002143.

Rouchy, J. M., and J. P. Saint-Martin. 1992. Late Miocene events in the Mediterranean as recorded by carbonate-evaporite relations. Geology 20:629–632.

Rutschmann, F., T. Eriksson, K. Abu Salim, and E. Conti. 2007. Assessing calibration uncertainty in molecular dating: the assignment of fossils to alternative calibration points. Syst. Biol. 56:591–608.

Salvo, G., G. Bacchetta, F. Ghahremaninejad, and E. Conti. 2008. Phylogenetic relationships of Ruteae (Rutaceae): new evidence from the chloroplast genome and comparisons with non-molecular data. Mol. Phylogenet. Evol. 49:736–748.

Sanderson, M. J., and J. A. Doyle. 2001. Sources of error and confidence intervals in estimating the age of angiosperms from *rbcL* and 18S rDNA data. Am. J. Bot. 88:1499–1516.

Sang, T., D. J. Crawford, and T. F. Stuessy. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). Am. J. Bot. 84:1120–1136.

Sanmartín, I. 2003. Dispersal vs. vicariance in the Mediterranean: historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). J. Biogeogr. 30:1883–1897.

Sanmartín, I., P. van der Mark, and F. Ronquist. 2008. Inferring dispersal: a Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. J. Biogeogr. 35:428–449.

Scotese, C. R. 2001. Atlas of Earth History, Volume 1, Paleogeography, PALEOMAP Project, Arlington, Texas.

Scott, K. D., C. L. McIntyre, and J. Playford. 2000. Molecular analyses suggest a need for a significant rearrangement of Rutaceae subfamilies and a minor reassessment of species relationships within *Flindersia*. Plant Syst. Evol. 223:15–27.

Selli, R. 1985. Tectonic evolution of the Tyrrhenian Sea*in* Geological evolution of the Mediterranean basin (D. J. Stanley, and F. C. Wezel, eds.). Springer-Verlag, New York.

Smith, H. V. 1938. Some new and interesting late Tertiary plants from Sucker Creek, Idaho-Oregon boundary. J. Torrey Bot. Soc. 65:557–564.

Smith, S. A. 2009. Taking into account phylogenetic and divergence-time uncertainty in a parametric biogeographical analysis of the Northern Hemisphere plant clade Caprifolieae. J. Biogeogr. 36:2324–2337.

Speranza, F., I. M. Villa, L. Sagnotti, F. Florindo, C. Cosentino, P. Cipollari, and M. Mattei. 2002. Age of the Corsica-Sardinia rotation and Liguro-Provençal basin spreading: new paleomagnetic and Ar/Ar evidence. Tectonophysics 347:231–251.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690.

Stampfli, G., J. Marcoux, and A. Baud. 1991. Tethyan margins in space and time. Palaeogeogr. Palaeoclimatol. Palaeoecol. 87:373–409.

Stebbins, G. L., and J. Major. 1965. Endemism and speciation in the California flora. Ecol. Monogr. 35:1–35.

Steininger, F. F., G. Rabeder, and F. Rögl. 1985. Land mammal distribution in the Mediterranean Neogene: a consequence of geokinematic and climatic events. Pages 559–571 *in* Geological evolution of the Mediterranean basin (D. J. Stanley, and F. C. Wezel, eds.). Springer-Verlag, New York.

Suc, J.-P. 1984. Origin and evolution of the Mediterranean vegetation and climate in Europe. Nature 307:429–432.

Sunding, P. 1979. Origins of the Macaronesian flora. Pages 13–40 *in* Plants and islands. (D. Bramwell, ed.) Academic Press, London.

Swofford, D. L. 2001. PAUP*4.0b10: Phylogenetic Analysis Using Parsimony (*and other methods). Sinauer, Sunderland, MA, USA.

Taberlet, P., L. Gielly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of 3 non-coding regions of chloroplast DNA. Plant Mol. Biol. 17:1105–1109.

Thompson, J. D. 2005. Plant evolution in the Mediterranean. Oxford University Press Inc., New York.

Thorne, J. L., and H. Kishino. 2002. Divergence time and evolutionary rate estimation with multilocus data. Syst. Biol. 51:689–702.

Thorne, J. L., H. Kishino, and I. S. Painter. 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol. 15.

Tiffney, B. H. 1980. Fruits and seeds of the Brandon Lignite, V. Rutaceae. J. Arnold Arboretum 61:1–40.

Tiffney, B. H. 1985. The Eocene North Atlantic Land Bridge: its importance in Tertiary and modern phytogeography of the Northern Hemisphere. J. Arnold Arboretum 66:243–273.

Tiffney, B. H. 1994. Re-evaluation of the age of the Brandon Lignite (Vermont, USA) based on plant megafossils. Rev. Palaeobot. Palyno. 82:299–315.

Townsend, C. C. 1968. Rutaceae. Pages 227–230 *in* Flora Europea, vol. 2, Rosaceae to Umbelliferae (T. G. Tutin, ed.) Cambridge University Press, Cambridge.

Upchurch, P. 2008. Gondwanan break-up: legacies of a lost world? Trends Ecol. Evol. 23:229–236.

Van Dam, J. A. 2006. Geographic and temporal patterns in the late Neogene (12-3 MA)

aridification of Europe: the use of small mammals as paleoprecipitation proxies. Palaeogeogr. Palaeoclimatol. Palaeoecol. 238:190–218.

Whittaker, R. J., and J. M. Fernandez-Palacios. 2007. Island biogeography: ecology, evolution, and conservation; 2nd ed, 2nd edition. Oxford University Press Inc., New York.

Whittaker, R. J., K. A. Triantis, and R. J. Ladle. 2008. A general dynamic theory of ocean island biogeography. J. Biogeogr. 35:977–994.

Wilgenbusch, J. C., D. L. Warren, and D. L. Swofford. 2004. AWTY: A system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. http://ceb.csit.fsu.edu/awty.

Wolfe, J. A. 1975. Some aspects of plant geography of the Northern Hemisphere during the Late Cretaceous and Tertiary. Ann. Mo. Bot. Gard. 62:264–279.

Wolfe, J. A. 1978. A paleobotanical interpretation of Tertiary climates in the Northern Hemisphere. Am. Sci. 66:694–703.

Yang, Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. Mol. Biol. Evol. 24:1586–1591.

Yesson, C., and A. Culham. 2006. Phyloclimatic modeling: combining phylogenetics and bioclimatic modeling. Syst. Biol. 55:785–802.

Yesson, C., N. H. Toomey, and A. Culham. 2009. *Cyclamen*: time, sea and speciation biogeography using a temporally calibrated phylogeny. J. Biogeogr. 36:1234–1252.

Young, N. D., and J. Healy. 2003. Gapcoder automates the use of indel characters in phylogenetic analysis. BMC Bioinformatics 4:6.

Figure legends

Figure 1. Distribution of *Ruta* and selected areas: (a) geographic distribution of the nine species of *Ruta*; (b) areas selected for the ancestral range reconstruction analyses, as defined by Mansion et al. (2008). Distribution of *R. corsica* and *R. lamarmorare* taken from Bacchetta et al. (2006).

Figure 2. Chronogram inferred with BEAST. Maximum-clade-credibility tree with mean nodal ages and 95% highest posterior density intervals indicated by grey bars. Nodes "a" to "f" indicate nodes of interest; nodes "1" to "4" indicate fossil constraints. Values next to branches represent Bayesian Posterior Probabilities / Maximum Parsimony Bootstrap Percentages / Maximum Likelihood Bootstrap Percentages. The grey rectangle shows the seven time slices ("S1" to "S7") used for the biogeographic analyses. The outgroup is not shown.

Figure 3. Biogeographic scenarios for *Ruta* and its Corso-Sardinian (C-S) endemics. On the left, dated phylogeny for *Ruta* and its sister group (taken from Fig. 2) showing: I, the origin (node a) and initial diversification (i.e., div.; node b) of *Ruta*; II, the origin of the C-S lineage (node c); III, the split between *R. corsica* (endemic to Corsica, C) and *R. lamarmorae* (endemic to Sardinia, S; node d). Vertical, coloured bars represent the time-windows of climatic/geologic events hypothesized to have affected the biogeography of *Ruta*: (1) beginning of a trend towards increasing aridification = 9-8 MA (Ivanov et al. 2002; Van Dam 2006); (2) onset of the Mediterranean climate = ~3-2 MA (Suc 1984; Thompson 2005); (3) split between the C-S microplate and the proto-Iberian peninsula = 30-28 MA (Alvarez et al. 1974; Rosenbaum et al. 2002a); (4) C-S block connected to, then disconnected from, the Apulian microplate = 20-9 MA (Cherchi and Montadert 1982; Rosenbaum et al. 2002a; Speranza et al. 2002; Rosenbaum and Lister 2004a); (5 and 7) Messinian Salinity Crisis = 5.96-5.33 MA (Gover et al. 2009); (6) formation of the Strait of Bonifacio = 15-9 MA (Alvarez et al. 1974; Cherchi and Montadert 1982); (8) Last Glacial Maximum = 0.02 MA (Lambeck et al. 2004; Lambeck and Purcell 2005). On the right, palaeo-maps associated with the nodes of interest, with corresponding letter. Maps A, B, and C were modified from Dercourt et al. (2000); map D was modified from Cavazza et al. (2004).

Figure 4. Biogeographic scenario for the Canarian clade of *Ruta*. Right, top: present distribution of the three Canarian endemics. Right, middle: maps of the Canary Islands associated with nodes e, f, and g, showing ancestral ranges reconstructed with Lagrange (reported also on the tree on the left; Table 1). Right, bottom: inferred, most-parsimonious number of dispersal events, indicated by arrows. Left, top: dated phylogeny for the Canarian endemics and their sister group; node "e": split between the Canarian endemics and *R. montana*; node "f": split between *R. oreojasme* and the rest; node "g": split between *R. microcarpa* and *R. pinnata*; node "h": crown node of *R. oreojasme*; node "i": crown node of *R. pinnata*. Left, bottom: gray rectangles: temporal envelopes for dispersal events 1 to 4, inferred along internodes connecting nodes e and f, f and h, f and g, g and i, respectively, spanning credibility intervals of nodes bracketing internodes (Table 1); colored rectangles: time intervals between island emergence (oldest estimate) and present time; hatched pattern: most likely time window for island colonization resulting from overlap between temporal envelopes for dispersal events and island emergence. Dashed orange line: onset of Mediterranean climate (~3 MA; Fernández-Palacios et al. 2008). Ages of the Canary Islands from Carracedo et al. (1998) and Anguita and Hernán (2000). MA = million years ago.

Figure 1

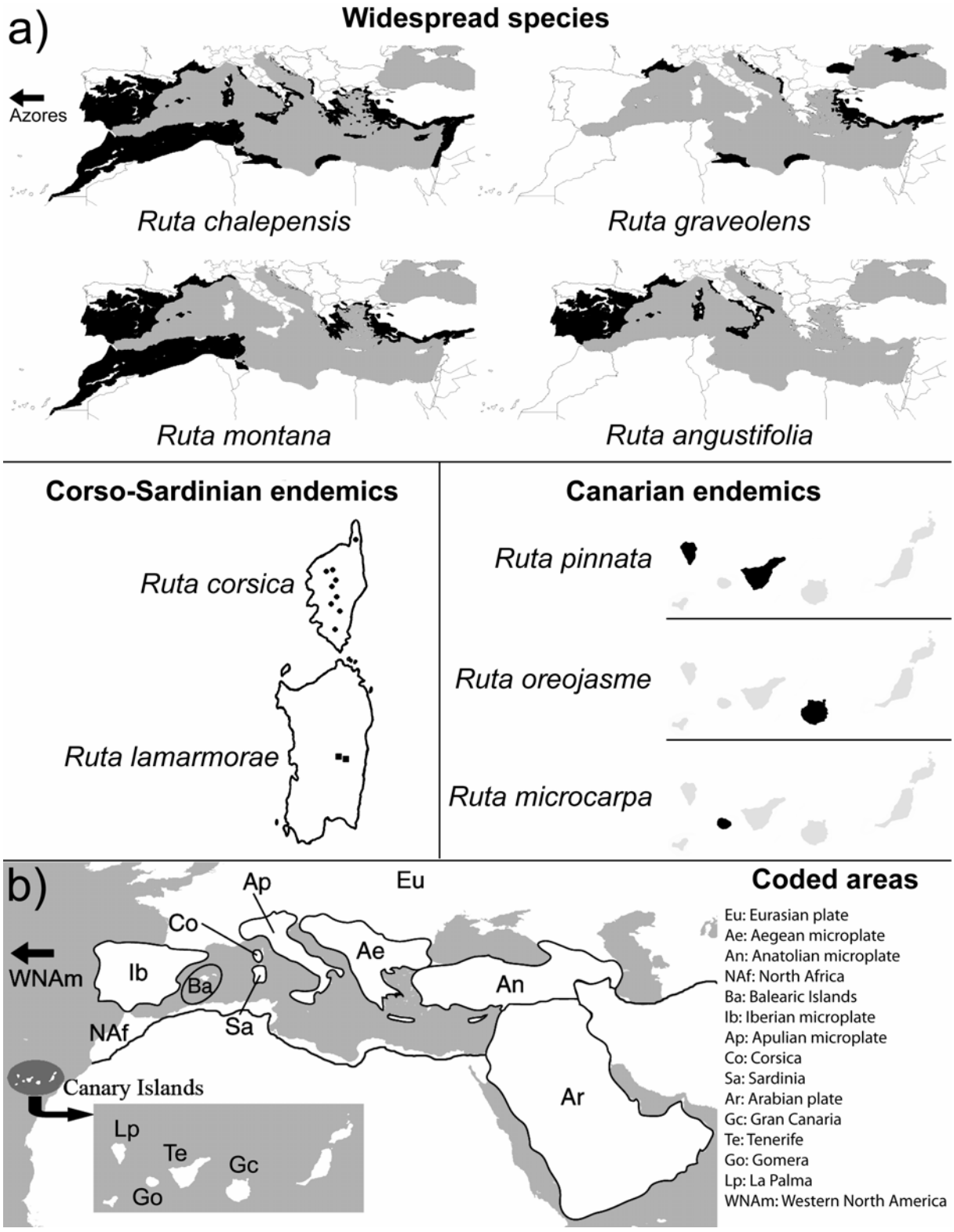

a)
Widespread species
Azores
Ruta chalepensis
Ruta graveolens
Ruta montana
Ruta angustifolia
Corso-Sardinian endemics
Ruta corsica
Ruta lamarmorae
Canarian endemics
Ruta pinnata
Ruta oreojasme
Ruta microcarpa
b)
Ap
Eu
Co
Ae
WNAm
Ib
Ba
An
NAf
Sa
Canary Islands
Ar
Lp
Te
Gc
Go
Coded areas
Eu: Eurasian plate
Ae: Aegean microplate
An: Anatolian microplate
NAf: North Africa
Ba: Balearic Islands
Ib: Iberian microplate
Ap: Apulian microplate
Co: Corsica
Sa: Sardinia
Ar: Arabian plate
Gc: Gran Canaria
Te: Tenerife
Go: Gomera
Lp: La Palma
WNAm: Western North America

Figure 2

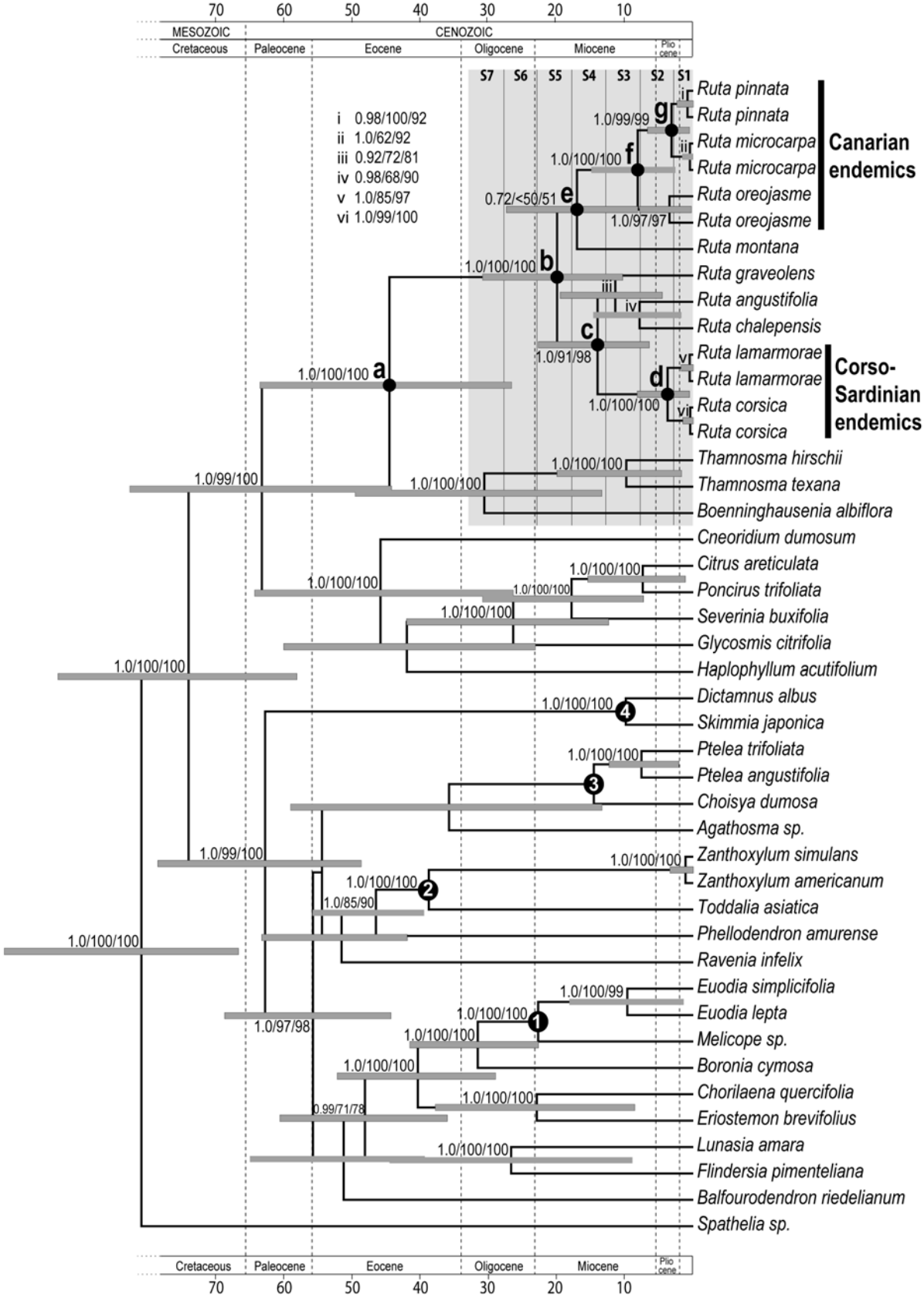

70
60
50
40
30
20
10
MESOZOIC
CENOZOIC
Cretaceous
Paleocene
Eocene
Oligocene
Miocene
Plio cene
S7
S6
S5
S4
S3
S2
S1
i 0.98/100/92
ii 1.0/62/92
iii 0.92/72/81
iv 0.98/68/90
v 1.0/85/97
vi 1.0/99/100
Ruta pinnata
Ruta pinnata
Ruta microcarpa
Ruta microcarpa
Ruta oreojasme
Ruta oreojasme
Canarian endemics
Ruta montana
Ruta graveolens
Ruta angustifolia
Ruta chalepensis
Ruta lamarmorae
Ruta lamarmorae
Ruta corsica
Ruta corsica
Corso-Sardinian endemics
Thamnosma hirschii
Thamnosma texana
Boenninghausenia albiflora
Cneoridium dumosum
Citrus areticulata
Poncirus trifoliata
Severinia buxifolia
Glycosmis citrifolia
Haplophyllum acutifolium
Dictamnus albus
Skimmia japonica
Ptelea trifoliata
Ptelea angustifolia
Choisya dumosa
Agathosma sp.
Zanthoxylum simulans
Zanthoxylum americanum
Toddalia asiatica
Phellodendron amurense
Ravenia infelix
Euodia simplicifolia
Euodia lepta
Melicope sp.
Boronia cymosa
Chorilaena quercifolia
Eriostemon brevifolius
Lunasia amara
Flindersia pimenteliana
Balfourodendron riedelianum
Spathelia sp.
1.0/99/99
1.0/100/100
0.72/<50/51
1.0/97/97
1.0/100/100
1.0/91/98
1.0/100/100
1.0/99/100
1.0/100/100
1.0/100/100
1.0/100/100
1.0/100/100
1.0/100/100
1.0/100/100
1.0/100/100
1.0/99/100
1.0/100/100
1.0/100/100
1.0/85/90
1.0/100/100
1.0/100/99
1.0/100/100
1.0/97/98
1.0/100/100
1.0/100/100
1.0/100/100
0.99/71/78
1.0/100/100
a
b
c
d
e
f
g
1
2
3
4

Figure 3

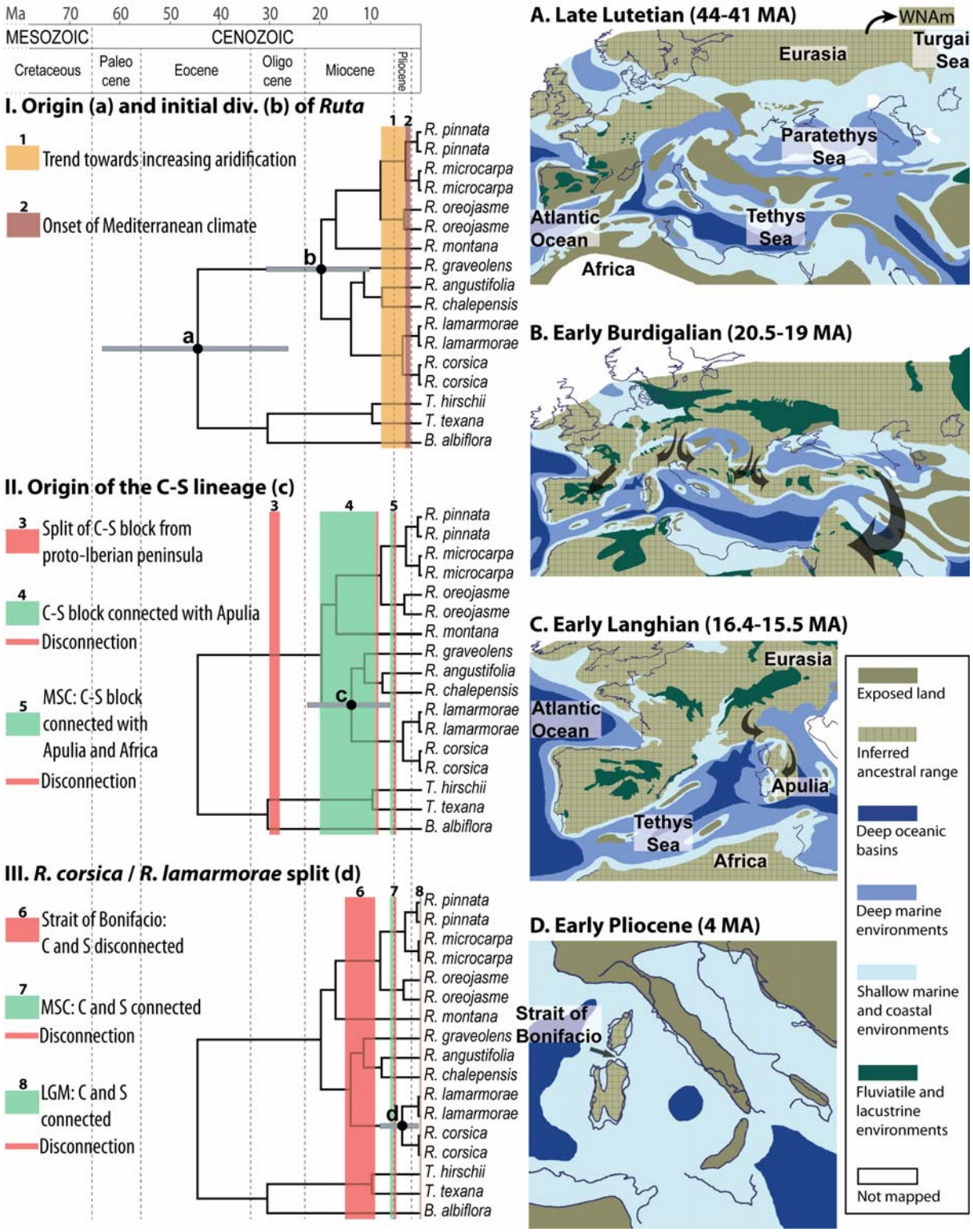

Ma 70 60 50 40 30 20 10
MESOZOIC
CENOZOIC
Cretaceous
Paleo cene
Eocene
Oligo cene
Miocene
Pliocene
I. Origin (a) and initial div. (b) of Ruta
1 Trend towards increasing aridification
2 Onset of Mediterranean climate
a
b
R. pinnata
R. pinnata
R. microcarpa
R. microcarpa
R. oreojasme
R. oreojasme
R. montana
R. graveolens
R. angustifolia
R. chalepensis
R. lamarmorae
R. lamarmorae
R. corsica
R. corsica
T. hirschii
T. texana
B. albiflora
II. Origin of the C-S lineage (c)
3 Split of C-S block from proto-Iberian peninsula
4 C-S block connected with Apulia
Disconnection
5 MSC: C-S block connected with Apulia and Africa
Disconnection
c
III. R. corsica / R. lamarmorae split (d)
6 Strait of Bonifacio: C and S disconnected
7 MSC: C and S connected
Disconnection
8 LGM: C and S connected
Disconnection
d
A. Late Lutetian (44-41 MA)
WNAm
Turgai Sea
Eurasia
Paratethys Sea
Atlantic Ocean
Tethys Sea
Africa
B. Early Burdigalian (20.5-19 MA)
C. Early Langhian (16.4-15.5 MA)
Eurasia
Atlantic Ocean
Apulia
Tethys Sea
Africa
D. Early Pliocene (4 MA)
Strait of Bonifacio
Exposed land
Inferred ancestral range
Deep oceanic basins
Deep marine environments
Shallow marine and coastal environments
Fluviatile and lacustrine environments
Not mapped

Figure 4

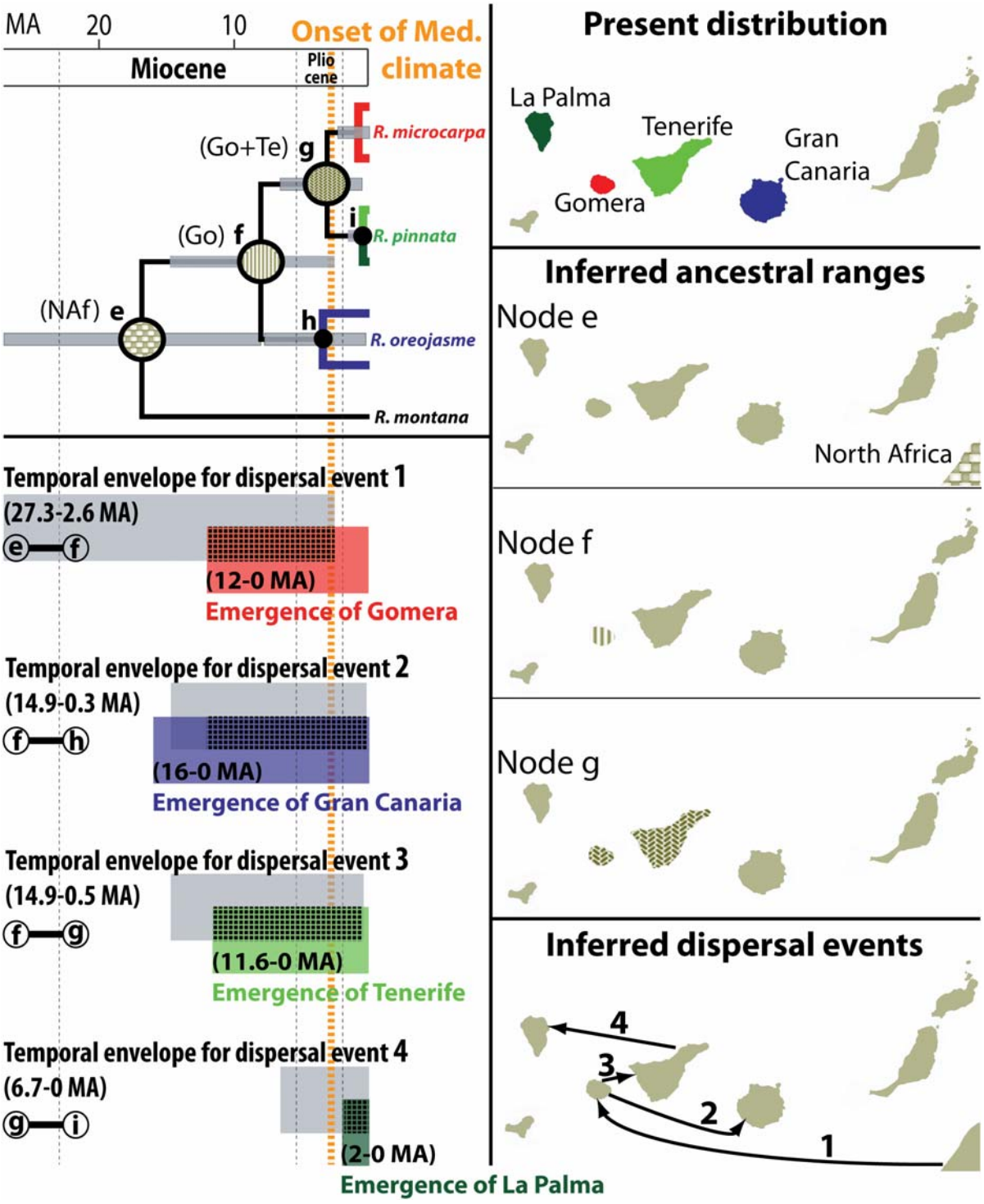

MA
20
10
Miocene
Plio cene
Onset of Med. climate
R. microcarpa
(Go+Te) g
i
R. pinnata
(Go) f
(NAf) e
h
R. oreojasme
R. montana
Temporal envelope for dispersal event 1
(27.3-2.6 MA)
e f
(12-0 MA)
Emergence of Gomera
Temporal envelope for dispersal event 2
(14.9-0.3 MA)
f h
(16-0 MA)
Emergence of Gran Canaria
Temporal envelope for dispersal event 3
(14.9-0.5 MA)
f g
(11.6-0 MA)
Emergence of Tenerife
Temporal envelope for dispersal event 4
(6.7-0 MA)
g i
(2-0 MA)
Emergence of La Palma
Present distribution
La Palma
Tenerife
Gran Canaria
Gomera
Inferred ancestral ranges
Node e
North Africa
Node f
Node g
Inferred dispersal events
4
3
2
1

Table 1. Results of molecular dating and ancestral range reconstruction analyses. The nodes of interest refer to Figure 2. CI = confidence interval: for BEAST this refers to the 95% highest posterior density intervals, for MULTIDIVTIME to the 95% credibility intervals. For area abbreviations see Figure 1.

		BEAST analysis	**MULTIDIVTIME analysis**	**Lagrange analysis with BEAST chronogram**	**Lagrange analysis with MULTIDIVTIME chronograms**
Node of interest	**Description**	**Mean nodal age (CI)**	**Mean nodal age (CI)**	**Ancestral range**	**Ancestral range**
a	Origin of *Ruta*	44.5596 (26.6484 - 63.5337)	46.2206 (32.0645 - 62.0346)	Eu + WNAm	Eu
b	Initial diversification of *Ruta*	19.9653 (10.3709 - 30.8809)	18.4885 (8.9494 - 32.3782)	Eu + Ae + An + Ba + Ib + Ap + Co + Sa + Ar + NAf	Eu + Ae + An + Ba + Ib + Ap + Co + Sa + Ar + NAf
c	Origin of the C-S lineage	14.0183 (6.4570 - 22.6956)	15.1908 (7.0254 - 27.6679)	Eu + Ae + An + Ba + Ib + Ap + Co + Sa + Ar + NAf	Co
d	Split between *R. corsica* and *R. lamarmorae*	3.7575 (0.5616 - 8.2029)	3.0826 (0.4987 - 8.6193)	Sa + Co	Sa + Co
e	Origin of the Canarian endemics	17.0426 (8.1353 - 27.3673)	16.8132 (7.9005 - 29.9693)	NAf	NAf
f	Split between *R. oreojasme* and the rest	8.1351 (2.6191 - 14.9420)	6.9065 (2.2322 - 15.7054)	Go	Go
g	Split between *R. pinnata* and *R. microcarpa*	3.1749 (0.5541 - 6.6831)	2.7843 (0.4336 - 8.1470)	Te + Go	Te + Go

Appendix 1. Taxa, sources, voucher information, and GenBank accession numbers. Voucher specimens are deposited in the following herbaria: Z = University of Zürich; K = Royal Botanic Gardens, Kew; MO = Missouri Botanical Garden; LE = V. L. Komarov Botanical Institute; NY = New York Botanical Garden.

Species name	**Source**	**Voucher**	***mat*K**	***rpl*16**	***trn*L-*trn*F**
RUTACEAE					
Ruta graveolens L.	France: Aveyron, Gages	Renaux 12 (Z)	EF489056	EF489130	EF489204
Ruta chalepensis L.	Italy: Rio di Pula, Villa San Pietro, Calgliari, Sardinia	Bacchetta 50 (Z)	EF489044	EF489118	EF489192
Ruta lamarmorae Bacch., Brullo & Giusso	Italy: Broncu Spina, Fonni, Nuoro, Sardinia	Bacchetta 162 (Z)	EF489050	EF489124	EF489198
Ruta lamarmorae Bacch., Brullo & Giusso	Italy: Bruncu Orisa, Villagrande Strisaili, Nuoro, Sardinia	Bacchetta & Carta 206 (Z)	EF489051	EF489125	EF489199
Ruta corsica DC.	France: Col deVergio, Corsica	Renaux 185 (Z)	EF489052	EF489126	EF489200
Ruta corsica DC.	France: Asco, Corsica	Renaux 198 (Z)	EF489053	EF489127	EF489201
Ruta angustifolia Pers.	France : Gard, St-Hippolyte du Fort	Renaux 264 (Z)	EF489058	EF489132	EF489206
Ruta montana (L.) L.	France: La Gardiole	Renaux 349 (Z)	EF489062	EF489136	EF489210
Ruta pinnata L.f.	Spain: Tenerife, Canary Is.	MA 655741	EF489065	EF489139	EF489213
Ruta pinnata L.f.	Spain: La Palma, Canary Is.	A. Santos 1.2007 (Z)	FJ716726	FJ716748	FJ716770
Ruta oreojasme Webb & Berth.	Spain: Gran Canaria, Bco Tirajuana, Canary Is.	Alfredo Amador 3 (Z)	EF489069	EF489143	EF489217
Ruta oreojasme Webb & Berth.	Spain: Barranco de Fataga, Gran Canaria, Canary Is.	GS_da JF1 (Z)	FJ716727	FJ716749	FJ716771
Ruta microcarpa Svent.	Spain: La Gomera, Canary Is.	P. Vargas 20/10/05 (Z)	EF489068	EF489142	EF489216
Ruta microcarpa Svent	Spain: La Abejera, Gomera, Canary Is.	GS_da JF2 (Z)	FJ716728	FJ716750	FJ716772
Boenninghausenia albiflora Reichb. ex Meissner	Japan	Chase 22071 (K)	EF489070	EF489144	EF489218

Dictamnus albus L.	Bosnia and Herzegovina	M. Bandle & A. Lenherr 22483 (Z)	EF489109	EF489183	EF489257
Cneoridium dumosum Hook.f.	USA: Oak Crest Park, California	Alexander Kocyan 154 (Z)	EF489108	EF489182	EF489256
Thamnosma hirschii Schweinf.	Yemen: Hadibu, Socotra	Mike Thiv 3187 (Z)	EF489071	EF489145	EF489219
Thamnosma texanum (A.Gray) Torrey	USA	M. Merello & D. Kruger 2643 (MO)	EF489075	EF489149	EF489223
Haplophyllum acutifolium (DC.) G.Don.	Iran	F. Ghahremaninejad 1469 (Z)	EF489076	EF489150	EF489224
Calodendrum capense (L.f.) Thunb.	Switzerland: Zürich Botanic Gardens; living collection	Cult. 20010674	EF489102	EF489176	EF489250
Zanthoxylum simulans Hance	Switzerland: Zürich Botanic Gardens; living collection	Cult. 19963815	EF489100	EF489174	EF489248
Zanthoxylum americanum P.Mill.	Spain: Real Jardin Botanico of Madrid; living collection	Gonzalo Nieto Feliner 4/22/05	EF489101	EF489175	EF489249
Choisya ternata Kunth	Switzerland: Zürich Botanic Gardens; living collection	Cult. 19963167	EF489104	EF489178	EF489252
Melicope sp.	France: New Caledonia	J. Munzinger & G. McPherson 785 (MO)	EF489107	EF489181	EF489255
Skimmia japonica Thunb.	Switzerland: Zürich Botanic Gardens; living collection	Cult. 19651137	EF489103	EF489177	EF489251
Orixa japonica Thunb.	Belgium: National Botanic Garden of Belgium; living collection	Cult. 19850295	EF489106	EF489180	EF489254
Citrus reticulata Blanco	Switzerland: Zürich Botanic Gardens; living collection	Cult. 19790418	FJ716729	FJ716751	FJ716773
Poncirus trifoliata (L.) Raf.		Chase 1767 (K)	FJ716730	FJ716752	FJ716774
Severinia buxifolia Ten.		Chase 1763 (K)	FJ716731	FJ716753	FJ716775
Glycosmis citrifolia Lindl.	Taiwan	Yih-Han Cahng 3310 (Z)	FJ716732	FJ716754	FJ716776
Euodia simplicifolia Ridl.		Chow and Wan 80125 (LE)	FJ716733	FJ716755	FJ716777
Euodia lepta (Spreng.) Merr.		Shin Ying Hu 11982 (LE)	FJ716734	FJ716756	FJ716778

Ptelea trifoliata L. var. *mollis* Torr. & A.Gray	Belgium: National Botanic Garden of Belgium; living collection	Cult. 19871850	FJ716735	FJ716757	FJ716779
Ptelea angustifolia Benth.		Chase 1765 (K)	FJ716736	FJ716758	FJ716780
Phellodendron amurense Rupr.		Chase 1771 (K)	FJ716737	FJ716759	FJ716781
Toddalia asiatica Baill.	Taiwan	K.U.Kramer, E.Zogg, H.Gassner 7828 (Z)	FJ716738	FJ716760	FJ716782
Spathelia sp.	UK: Kew DNA bank 1899		FJ716739	FJ716761	FJ716783
Lunasia amara Blanco		Chase 1347 (K)	FJ716740	FJ716762	FJ716784
Flindersia pimenteliana F. Muell.	Australia	PIF 29742 (Z)	FJ716741	FJ716763	FJ716785
Chorilaena quercifolia Endl.		Chase 1755 (K)	FJ716742	FJ716764	FJ716786
Eriostemon brevifolius A. Cunn. ex Endl.	UK: Kew DNA bank 2061		FJ716743	FJ716765	FJ716787
Boronia cymosa Endl.		Chase 2186 (K)	FJ716744	FJ716766	FJ716788
Agathosma sp.	UK: Kew DNA bank 2342		FJ716745	FJ716767	FJ716789
Ravenia infelix Vell.		J. Kallunki *et al*. 614 (NY)	FJ716746	FJ716768	FJ716790
Balfourodendron riedelianum (Engl.) Engl.	UK: Kew DNA bank 2345		FJ716747	FJ716769	FJ716791
SIMAROUBACEAE					
Ailanthus giraldii Dode		Chase 16984 (K)	EF489112	EF489186	EF489260
Simarouba glauca DC.		Chase 124 (K)	EF489113	EF489187	EF489261
MELIACEAE					
Cipadessa baccifera (Roth) Miq.	Switzerland: Zürich Botanic Gardens; living collection	Cult. 20011380	EF489116	EF489190	EF489264
Melia azedarach L.	Switzerland: Zürich Botanic Gardens; living collection	Cult. 20012236	EF489117	EF489191	EF489265

Appendix 2. Detailed information on the four fossils of Rutaceae used as calibration points for the molecular dating analyses. Absolute stratigraphic dates were taken from Gradstein *et al*. (2004).

Fossil 1: *Euodia lignita* Tiffney.
Material: Seeds.
Locality: Brandon Lignite, Forestdale, Massachussetts, USA.
Age: Early Oligocene to Early Miocene (33.9-15.97 MA).
Affinities with extant taxa: *Euodia*.
Diagnosis: "The seeds average 4.5 mm long and 2.8 mm in diameter. [...] All are marked by a hilar scar that extends from the apex to the base of the seed and is bordered by a wide margin. At the base of the hilum a short raphe leads to the large, pitlike basal chalaza. The micropyle is at the apex of the seed, just beyond the terminus of the hilum. The dull black external surface of the sclerotesta is marked by many faint longitudinal ridges that are crossed at intervals by weak, short transverse ridges [...]. The inner layer of the outer integument (sclerotesta) is 150-250 μm thick and is formed of many layers of isodiametric sclereids" [...] (Tiffney 1980; p.7).
Remarks: "No single modern species is completely similar to the fossil, and only one that has been examined, *Euodia colorata*, is as large. [...] Thus, although the fossil is similar to certain modern species, it is not completely comparable to any one, and it is best to regard it as belonging to an extinct species" (Tiffney 1980; p.7).
Placement of fossil in phylogeny: Stem node of *Euodia* (node 1 in Fig. 2).
References: Tiffney (1980, 1994).

Fossil 2: *Toddalia excavata* (Chandler) Gregor**.**
Material: Seed.
Locality: Cliff End, Mudeford, England, UK.
Age: Uppermost Auversian (Bartonian; 40.4-37.2 MA).
Affinities with extant taxa: *Toddalia*.
Diagnosis: "Small reniform seed, bisymmetric, with deeply excavated triangular hilar scar. Finely pitted surface with equiaxial cells (0.017mm). Obscure concentric ridges parallel to the dorsal margin" (Gregor 1979; p.321).
Remarks: "The very small seed undoubtedly belongs to the genus *Toddalia*, as I have seen myself in the British Museum" (Gregor 1979; p.321).
Placement of fossil in phylogeny: Stem node of *Toddalia* (node 2 in Fig. 2).
References: Gregor (1979).

Fossil 3: *Ptelea enervosa* H.V. Smith.
Material: Samaras.
Locality: Ballantyne Ranch, Succor Creek valley, Idaho-Oregon boundary, USA.
Age: Middle Miocene (15.97-11.608 MA).
Affinities with extant taxa: *Ptelea*.
Diagnosis: "The holotype of *Ptelea enervosa* is a faintly preserved impression that has been coated with a clear preservative. Despite these impediments to recognizing distinguishing characters on the fossil, certain characters consistent with those of extant *Ptelea* samaras are preserved, including the impression of the floral disk at the base of the samara; samara wing formed from two semicircular wings completely fused to one another at the base and apex; presence of more than one vein extending medially from the disk to the central body; and single vein extending from the apex of the fruit body to the samara apex" (Call and Dilcher 1995; p.1072).
Remarks: "*Ptelea enervosa* samaras are the only indisputable fossil fruits of the extant rutaceous genus *Ptelea*" (Call and Dilcher 1995; p.1073).

Placement of fossil in phylogeny: Stem node of *Ptelea* (node 3 in Fig. 2).
References: Smith (1938), Call and Dilcher (1995).

Fossil 4: *Skimmia tortonica* Palamarev & Usunova**.**
Material: Leaves.
Locality: Milčina lâka, Bezirk Vidin, Bulgaria.
Age: Tortonian (11.608-7.246 MA).
Affinities with extant taxa: *Skimmia*.
Diagnosis: Leathery, lanceolate, entire leaf, 11 cm long, 3 cm broad, on both sides 9 nerves with angles of 45-55°. Upper epidermis polygonal. [...] [translated from German]
Remarks: "Stark ausgeprägt ist insbesondere seine Beziehung zu Skimmia und am meisten zu der gegenwärtigen Art S. japonica Thunb. Die Morphologie und Anatomie dieser Art stimmen mit dem fossilen Taxon fast ganz überein. Der Unterschied besteht nur in den gewellten Zellwänden bei der rezenten Art" (Palamarev and Usunova 1970; p.836).
Placement of fossil in phylogeny: Crown node of *Skimmia* (node 4 in Fig. 2).
References: Palamarev and Usunova (1970).

Appendix 3. Setup and results of the sensitivity analyses, implemented in MULTIDIVTIME and BEAST, carried out to assess the impact of prior choice on resulting age estimates.

MULTIDIVTIME

Root priors:

1) rttm = 65 MA[1]; rttmsd = 60 MA; bigtime = 125 MA[2]
2) rttm = 65 MA[1]; rttmsd = 60 MA; bigtime = 90.5 MA[3]
3) rttm = 65 MA[1]; rttmsd = 30 MA; bigtime = 125 MA[2]

[1]Age of oldest reliable fossil of Rutaceae (Knobloch and Mai 1986).
[2]Appearance of tricolpate pollen in the fossil record (Sanderson and Doyle 2001; Anderson et al. 2005).
[3]Age of Sapindales estimated by Muellner et al. (2007).

Tree priors:

1) 50% majority rule consensus tree inferred with MRBAYES (not fully resolved).
2) MAP tree inferred with MRBAYES (fully resolved).

	Root prior 1	**Root prior 2**	**Root prior 3**
Nodes of interest	**Tree prior 1**	**Tree prior 1**	**Tree prior 1**
Node a	46.2206 (32.0645-62.0346)	38.8814 (27.6945-50.5887)	45.2930 (31.3828-60.5606)
Node b	18.4885 (8.9494-32.3782)	16.7216 (7.8082-29.6077)	18.0692 (8.8101-31.4944)
Node c	15.1908 (7.0254-27.6679)	13.7784 (6.2193-25.6406)	14.8361 (6.9803-26.8048)
Node d	3.0826 (0.4987-8.6193)	2.8731 (0.4793-8.2064)	2.9968 (0.4859-8.0625)
Node e	16.8132 (7.9005-29.9693)	15.2737 (6.9984-27.6312)	16.4395 (7.7951-29.1296)
Node f	6.9065 (2.2322-15.7054)	6.4586 (2.0037-14.9094)	6.7445 (2.2384-15.1033)
Root	113.0689 (91.6502-124.5194)	85.6628 (75.5795-90.3309)	111.2952 (89.0425-124.3514)
	Root prior 1	**Root prior 2**	**Root prior 3**
Nodes of interest	**Tree prior 2**	**Tree prior 2**	**Tree prior 2**
Node a	48.3096 (33.8692-64.0699)	39.7181 (28.9733-50.9774)	47.4969 (33.1730-62.8971)
Node b	19.4986 (9.4506-33.9525)	17.1426 (8.4592-29.2790)	19.3551 (9.3625-33.8412)
Node c	15.9865 (7.4480-28.9449)	14.0952 (6.4987-25.3462)	15.9105 (7.3756-28.9178)
Node d	3.2736 (0.5603-9.2817)	2.9791 (0.4913-8.3564)	3.3000 (0.5260-9.2712)
Node e	17.7433 (8.4758-31.5835)	15.6665 (7.4683-27.5476)	17.6318 (8.3472-31.4751)
Node f	7.3294 (2.4124-16.4594)	6.6211 (2.0703-14.9698)	7.3027 (2.3349-16.5934)
Root	114.1278 (93.7951-124.5844)	86.1116 (76.5002-90.3549)	112.5363 (91.6001-124.4794)

BEAST

Trials:

1) Lognormal prior distributions for fossil calibrations[1]

2) Lognormal prior distributions for fossil calibrations + root prior[2]

3) Different lognormal prior distributions for fossil calibrations[3]

4) Different lognormal prior distributions for fossil calibrations + root prior

5) Lognormal prior distributions for fossil calibrations + different root prior[4]

[1]95% of the prior distribution falling between the geological interval of the fossil in question.

[2]Root prior modelled with a lognormal distribution corresponding to the oldest reliable fossil of Rutaceae (Knobloch and Mai 1986).

[3]Lognormal prior distributions for fossil calibrations more dispersed than before.

[4]Root prior modelled with a lognormal distribution corresponding to the age of Sapindales estimated by Muellner et al. (2007).

Nodes of interest	Trial 1	Trial 2	Trial 3
Node a	52.894 (26.3949-84.3561)	44.5596 (26.6484-63.5337)	53.3379 (28.5726-80.3945)
Node b	23.818 (10.7687-39.6131)	19.9653 (10.3709-30.8809)	23.0419 (11.0187-38.1719)
Node c	16.5512 (6.625-28.5061)	14.0183 (6.4570-22.6956)	16.3329 (6.7596-27.8501)
Node d	4.6094 (0.6268-10.1325)	3.7575 (0.5616-8.2029)	4.2406 (0.6414-9.1804)
Node e	20.1509 (8.1082-34.2989)	17.0426 (8.1353-27.3673)	19.6835 (7.9741-33.1409)
Node f	9.8083 (3.0815-18.1869)	8.1351 (2.6191-14.9420)	9.2805 (3.0476-17.1166)
Root	109.5337 (64.3919-162.3759)	80.9321 (66.7178-101.0489)	110.6569 (64.2797-170.1478)

	Trial 4	Trial 5
Node a	44.8534 (27.4531-62.1783)	42.1964 (25.0373-58.0715)
Node b	19.0607 (9.7601-30.0505)	19.7349 (10.6908-30.3561)
Node c	13.5394 (6.3363-22.2425)	13.9631 (6.8098-22.3633)
Node d	3.5451 (0.5880-7.5786)	3.8495 (0.4901-8.1862)
Node e	16.3261 (8.0658-26.3728)	16.8191 (8.1193-26.5291)
Node f	7.6565 (2.6265-13.7786)	8.0817 (2.6061-14.1776)
Root	82.1821 (66.8807-103.4270)	78.4152 (71.6909-86.1987)

Appendix 4. Distribution of *Ruta* and sister group coded for the selected areas. Areas are shown in Fig. 1. 0 = absence; 1 = presence.

	Eu	**Ae**	**An**	**NAf**	**Ba**	**Ib**	**Ap**	**Co**	**Sa**	**Ar**	**Gc**	**Te**	**Lp**	**Go**
Ruta graveolens	1	1	1	1	0	0	0	0	0	0	0	0	0	0
Ruta montana	1	1	1	1	1	1	1	0	0	0	0	0	0	0
Ruta angustifolia	1	1	0	0	1	1	1	1	1	0	0	0	0	0
Ruta corsica	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Ruta lamarmorae	0	0	0	0	0	0	0	0	1	0	0	0	0	0
Ruta chalepensis	1	1	1	1	1	1	1	1	1	1	0	0	0	0
Ruta oreojasme	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Ruta pinnata	0	0	0	0	0	0	0	0	0	0	0	1	1	0
Ruta microcarpa	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Boenninghausenia japonica	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Thamnosma texana	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Thamnosma hirschii	0	0	0	1	0	0	0	0	0	1	0	0	0	0

	WNAm
Ruta graveolens	0
Ruta montana	0
Ruta angustifolia	0
Ruta corsica	0
Ruta lamarmorae	0
Ruta chalepensis	0
Ruta oreojasme	0
Ruta pinnata	0
Ruta microcarpa	0
Boenninghausenia japonica	0
Thamnosma texana	1
Thamnosma hirschii	0

Appendix 5. Matrices of scaling factors for the rate of dispersal between areas (Fig. 1), which were constructed for seven time slices ("S1" to "S7") and implemented in the ancestral range reconstruction analyses carried out with Lagrange.

S1
0 - 2.5 MA

	Eu	Ae	An	NAf	Ba	Ib	Ap	Co	Sa	Ar	Gc	Te	Lp	Go	WNAm
Eu	-														
Ae	1	-													
An	1	0.9	-												
NAf	0.5	0.6	0.6	-											
Ba	0.7	0.5	0.4	0.7	-										
Ib	1	0.5	0.4	0.9	0.8	-									
Ap	1	1	0.5	0.7	0.6	0.6	-								
Co	0.8	0.5	0.4	0.6	0.7	0.6	0.8	-							
Sa	0.7	0.5	0.4	0.7	0.7	0.6	0.8	0.9	-						
Ar	0.5	0.5	1	1	0.4	0.4	0.5	0.4	0.4	-					
Gc	0.3	0.3	0.3	0.8	0.4	0.5	0.3	0.4	0.4	0.3	-				
Te	0.3	0.3	0.3	0.7	0.4	0.5	0.3	0.4	0.4	0.3	0.9	-			
Lp	0.3	0.3	0.3	0.6	0.4	0.5	0.3	0.4	0.4	0.3	0.6	0.8	-		
Go	0.3	0.3	0.3	0.7	0.4	0.5	0.3	0.4	0.4	0.3	0.7	0.9	0.9	-	
WNAm	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	-

S2
2.5 - 7.5 MA

	Eu	Ae	An	NAf	Ba	Ib	Ap	Co	Sa	Ar	Gc	Te	Lp	Go	WNAm
Eu	-														
Ae	1	-													
An	1	0.9	-												
NAf	0.5	0.6	0.6	-											
Ba	0.7	0.5	0.4	0.7	-										
Ib	1	0.5	0.4	0.9	0.8	-									
Ap	0.9	1	0.5	0.7	0.6	0.6	-								
Co	0.8	0.5	0.4	0.6	0.7	0.6	0.8	-							
Sa	0.7	0.5	0.4	0.7	0.7	0.6	0.8	0.9	-						
Ar	0.5	0.5	1	1	0.4	0.4	0.5	0.4	0.4	-					
Gc	0.3	0.3	0.3	0.8	0.4	0.5	0.3	0.4	0.4	0.3	-				
Te	0.3	0.3	0.3	0.7	0.4	0.5	0.3	0.4	0.4	0.3	0.9	-			
Lp	0	0	0	0	0	0	0	0	0	0	0	0	-		
Go	0.3	0.3	0.3	0.7	0.4	0.5	0.3	0.4	0.4	0.3	0.7	0.9	0	-	
WNAm	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	-

S3

7.5 - 12.5 MA	Eu	Ae	An	NAf	Ba	Ib	Ap	Co	Sa	Ar	Gc	Te	Lp	Go	WNAm
Eu	-														
Ae	1	-													
An	1	0.9	-												
NAf	0.5	0.6	0.6	-											
Ba	0.7	0.5	0.4	0.7	-										
Ib	1	0.5	0.4	0.9	0.8	-									
Ap	0.9	1	0.5	0.7	0.6	0.6	-								
Co	0.8	0.5	0.4	0.6	0.7	0.6	1	-							
Sa	0.7	0.5	0.4	0.7	0.7	0.6	1	0.9	-						
Ar	0.5	0.5	0.9	1	0.4	0.4	0.5	0.4	0.4	-					
Gc	0.3	0.3	0.3	0.8	0.4	0.5	0.3	0.4	0.4	0.3	-				
Te	0.2	0.2	0.2	0.6	0.3	0.4	0.2	0.3	0.3	0.2	0.8	-			
Lp	0	0	0	0	0	0	0	0	0	0	0	0	-		
Go	0.2	0.2	0.2	0.6	0.3	0.4	0.2	0.3	0.3	0.2	0.6	0.8	0	-	
WNAm	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	-

S4

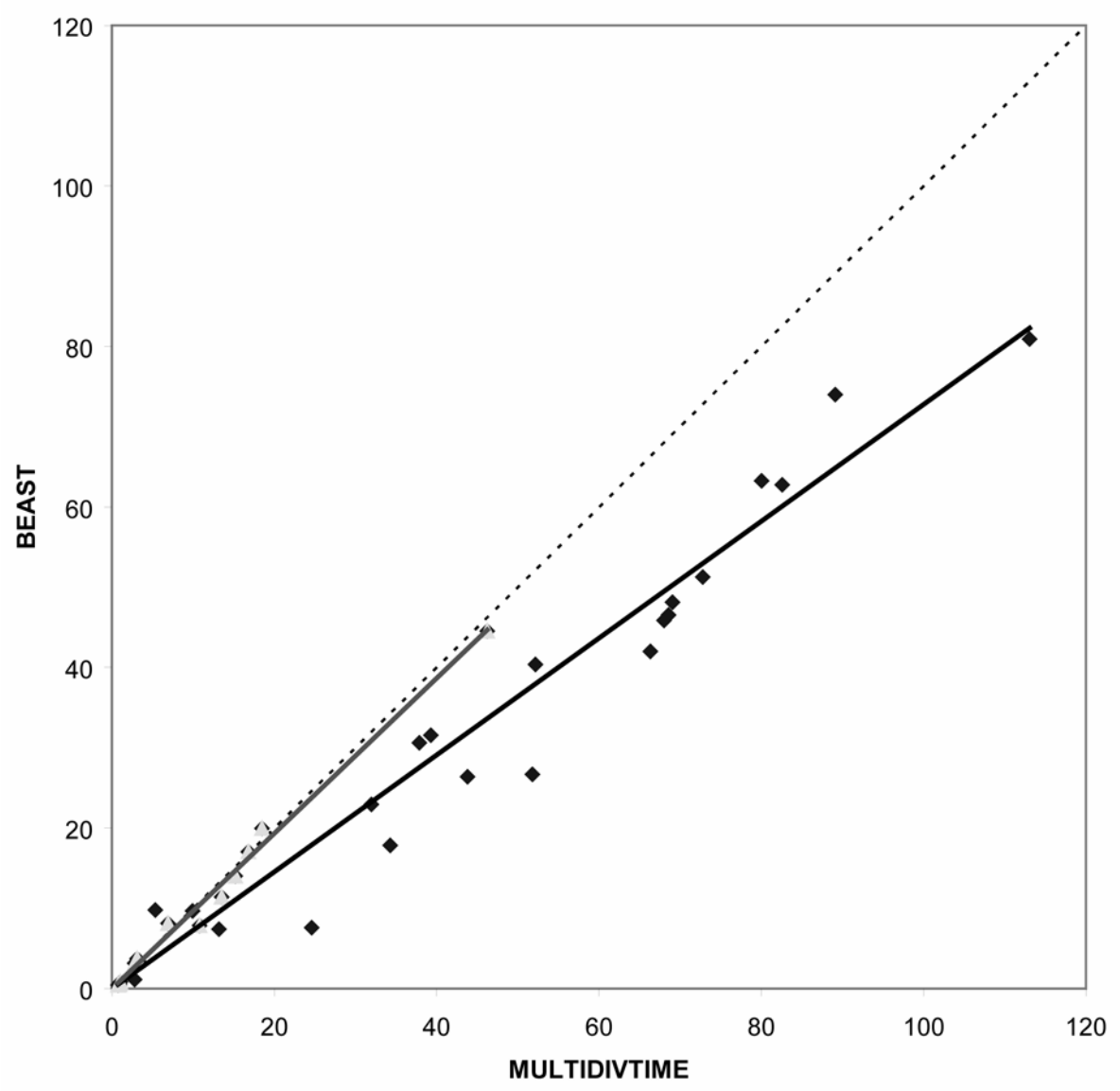
12.5 - 17.5 MA	Eu	Ae	An	NAf	Ba	Ib	Ap	Co	Sa	Ar	Gc	Te	Lp	Go	WNAm
Eu	-														
Ae	1	-													
An	1	0.9	-												
NAf	0.5	0.6	0.6	-											
Ba	0.7	0.5	0.4	0.7	-										
Ib	1	0.5	0.4	0.8	0.8	-									
Ap	0.9	1	0.5	0.7	0.6	0.6	-								
Co	0.8	0.5	0.4	0.6	0.7	0.6	1	-							
Sa	0.7	0.5	0.4	0.7	0.7	0.6	1	0.9	-						
Ar	0.5	0.5	0.8	1	0.4	0.4	0.5	0.4	0.4	-					
Gc	0.2	0.2	0.2	0.7	0.3	0.4	0.2	0.3	0.3	0.2	-				
Te	0	0	0	0	0	0	0	0	0	0	0	-			
Lp	0	0	0	0	0	0	0	0	0	0	0	0	-		
Go	0	0	0	0	0	0	0	0	0	0	0	0	0	-	
WNAm	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0	0	-

S5 **17.5 - 22.5 MA**	Eu	Ae	An	NAf	Ba	Ib	Ap	Co	Sa	Ar	Gc	Te	Lp	Go	WNAm
Eu	-														
Ae	1	-													
An	1	0.9	-												
NAf	0.5	0.6	0.6	-											
Ba	0.7	0.5	0.4	0.7	-										
Ib	1	0.5	0.4	0.8	0.8	-									
Ap	0.9	1	0.5	0.7	0.6	0.6	-								
Co	0.8	0.5	0.4	0.6	0.7	0.6	1	-							
Sa	0.7	0.5	0.4	0.7	0.7	0.6	0.9	1	-						
Ar	0.5	0.5	0.8	1	0.4	0.4	0.5	0.4	0.4	-					
Gc	0	0	0	0	0	0	0	0	0	0	-				
Te	0	0	0	0	0	0	0	0	0	0	0	-			
Lp	0	0	0	0	0	0	0	0	0	0	0	0	-		
Go	0	0	0	0	0	0	0	0	0	0	0	0	0	-	
WNAm	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0	0	0	-

S6 **22.5 - 27.5 MA**	Eu	Ae	An	NAf	Ba	Ib	Ap	Co	Sa	Ar	Gc	Te	Lp	Go	WNAm
Eu	-														
Ae	1	-													
An	1	0.9	-												
NAf	0.5	0.6	0.6	-											
Ba	0.8	0.5	0.4	0.7	-										
Ib	1	0.5	0.4	0.8	0.9	-									
Ap	0.9	1	0.5	0.7	0.6	0.6	-								
Co	0.9	0.5	0.4	0.6	0.7	0.7	1	-							
Sa	0.8	0.5	0.4	0.7	0.8	0.8	0.8	1	-						
Ar	0.5	0.5	0.8	1	0.4	0.4	0.5	0.4	0.4	-					
Gc	0	0	0	0	0	0	0	0	0	0	-				
Te	0	0	0	0	0	0	0	0	0	0	0	-			
Lp	0	0	0	0	0	0	0	0	0	0	0	0	-		
Go	0	0	0	0	0	0	0	0	0	0	0	0	0	-	
WNAm	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0	0	0	-

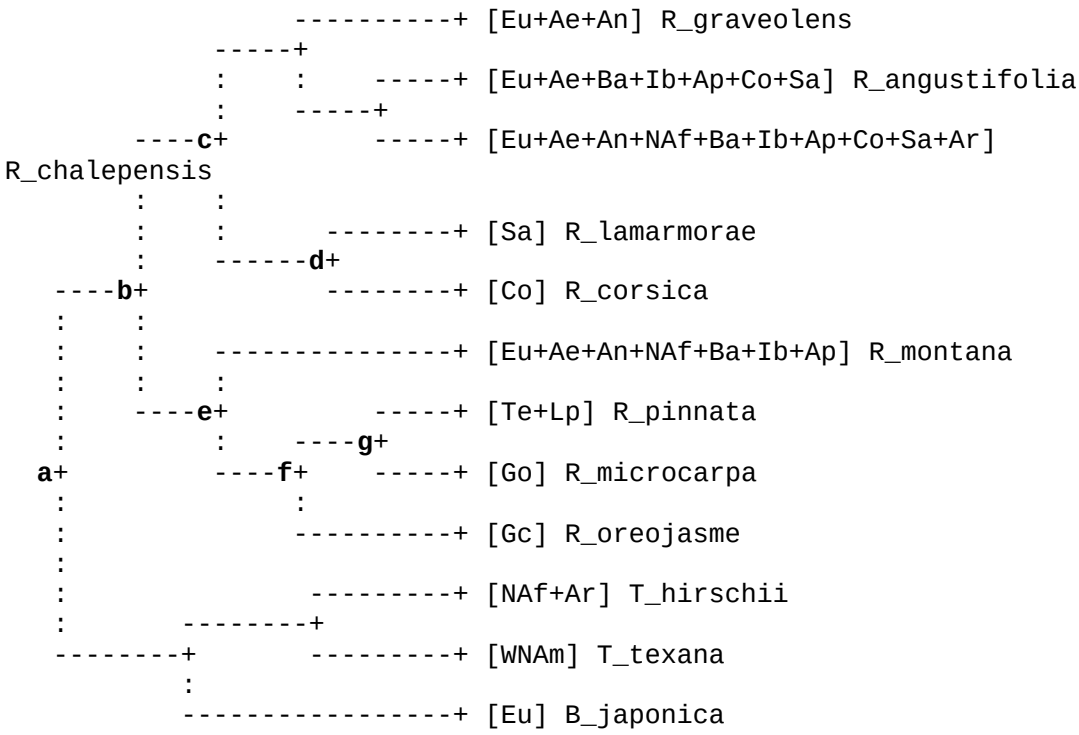
S7 27.5 - 32.5 MA	Eu	Ae	An	NAf	Ba	Ib	Ap	Co	Sa	Ar	Gc	Te	Lp	Go	WNAm
Eu	-														
Ae	0.8	-													
An	1	0.8	-												
NAf	0.4	0.6	0.5	-											
Ba	1	0.4	0.4	0.7	-										
Ib	1	0.4	0.4	0.8	1	-									
Ap	0.9	1	0.5	0.7	0.6	0.6	-								
Co	1	0.4	0.4	0.7	1	1	0.9	-							
Sa	1	0.4	0.4	0.7	1	1	0.7	1	-						
Ar	0.4	0.5	0.7	1	0.4	0.4	0.5	0.4	0.4	-					
Gc	0	0	0	0	0	0	0	0	0	0	-				
Te	0	0	0	0	0	0	0	0	0	0	0	-			
Lp	0	0	0	0	0	0	0	0	0	0	0	0	-		
Go	0	0	0	0	0	0	0	0	0	0	0	0	0	-	
WNAm	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0	0	0	-

Appendix 6. Correlation between age estimates obtained with two relaxed clock methods: the Bayesian uncorrelated lognormal method, implemented with BEAST, versus the Bayesian autocorrelation method, implemented with MULTIDIVTIME. The black and gray lines are linear regressions through the origin for all nodes of the tree (black diamonds), and for the nodes pertaining to *Ruta* and its sister group only (gray triangles), respectively. Strictly equal ages are indicated by the dotted line. Values in the axes represent million of years before present.



Appendix 7. Results of the Lagrange analyses.

Cladogram showing nodes of interest ("a" to "g") and geographic distribution at the tips:



Ancestral range subdivision/inheritance scenarios ("splits") for the nodes of interest.

* Split format: [left|right], where "left" and "right" are the ranges inherited by each descendant branch (on the tree above, "left" is the upper branch, and "right" the lower branch).

* Only splits within 1 log-likelihood unit of the maximum for each node are shown. "Rel.Prob" is the relative probability (fraction of the global likelihood) of a split.

1. Areas optimized onto the maximum-clade-credibility tree inferred with BEAST.

Global ML at root node: -lnL = 62.58 dispersal = 0.06274 extinction = 0.0335

At node a (origin of *Ruta*):

SPLIT	lnL	Rel.Prob
[Eu|Eu+WNAm]	**-64.96**	0.09256
[Eu|WNAm]	-65.06	0.08406

At node b (initial diversification of *Ruta*):

SPLIT	lnL	Rel.Prob
[Eu+Ae+An+Ba+Ib+Ap+Co+Sa+Ar|NAf]	**-64.41**	0.1603
[Eu+Ae+An+NAf+Ba+Ib+Ap+Co+Sa+Ar|NAf]	-64.42	0.1593

At node c (origin of the C-S lineage):

SPLIT	lnL	Rel.Prob

[Eu+Ae+An+NAf+Ba+Ib+Ap+Co+Ar|Sa]	**-65.07**	0.08304
[Eu+Ae+An+NAf+Ba+Ib+Ap+Co+Sa+Ar|Sa]	-65.11	0.07983
[Eu+Ae+An+NAf+Ba+Ib+Ap+Sa+Ar|Co]	-65.14	0.0771
[Eu+Ae+An+NAf+Ba+Ib+Ap+Co+Sa+Ar|Co]	-65.16	0.07543

At node d (split between *R. corsica* and *R. lamarmorae*):

SPLIT	lnL	Rel.Prob
[Sa|Co]	**-63.02**	0.642

At node e (origin of the Canarian endemics):

SPLIT	lnL	Rel.Prob
[NAf|NAf]	**-63.99**	0.2435
[NAf|Go]	-64.64	0.1271
[NAf|Te]	-64.88	0.1002

At node f (split of *R. oreojasme* from the rest):

SPLIT	lnL	Rel.Prob
[Go|Go]	**-64.66**	0.1245
[Go|Gc]	-64.79	0.1095
[Te|Te]	-64.94	0.09424
[Te|Gc]	-65.04	0.08516
[Lp|Lp]	-65.26	0.06834
[Lp|Gc]	-65.45	0.0568

At node g (split between *R. pinnata* and *R. microcarpa*):

SPLIT	lnL	Rel.Prob
[Te|Go]	**-63.57**	0.3729
[Lp|Go]	-63.71	0.3228

2. Areas optimized onto the MULTIDIVTIME chronogram, which was inferred by using the 50% majority rule consensus tree from MRBAYES as input tree.

Global ML at root node: -lnL = 64.8 dispersal = 0.0499 extinction = 0.0721

At node a (origin of *Ruta*):

SPLIT	lnL	Rel.Prob
[Eu|Eu]	**-70.5**	0.003341
[Eu|Eu+WNAm]	-70.64	0.002915
[An|An]	-70.73	0.00267
[Eu|WNAm]	-70.8	0.002488
[Ae|Ae]	-70.85	0.002348
[Ib|Ib]	-71.04	0.001944

At node b (initial diversification of *Ruta*):

SPLIT	lnL	Rel.Prob
[Eu+Ae+An+Ba+Ib+Ap+Co+Sa+Ar|NAf]	**-68.31**	0.02986
[Eu+Ae+An+NAf+Ba+Ib+Ap+Co+Sa+Ar|NAf]	-68.36	0.0286

At node c (origin of the C-S lineage):

SPLIT	lnL	Rel.Prob

[Co|Co]	**-67.49**	0.06783
[Ap|Ap]	-68.15	0.03494

At node d (split between *R. corsica* and *R. lamarmorae*):

SPLIT	lnL	Rel.Prob
[Sa|Co]	**-65.48**	0.5069

At node e (origin of the Canarian endemics):

SPLIT	lnL	Rel.Prob
[NAf|NAf]	**-65.87**	0.3418

At node f (split of *R. oreojasme* from the rest):

SPLIT	lnL	Rel.Prob
[Go|Go]	**-66.61**	0.1641
[Te|Te]	-66.78	0.1387
[Lp|Lp]	-67.16	0.09441
[Go|Gc]	-67.43	0.07234
[Te|Gc]	-67.57	0.06254

At node g (split between *R. pinnata* and *R. microcarpa*):

SPLIT	lnL	Rel.Prob
[Te|Go]	**-65.98**	0.3069
[Lp|Go]	-66.15	0.2584

3. Areas optimized onto the MULTIDIVTIME chronogram, which was inferred by using the MAP tree from MRBAYES as input tree.

Global ML at root node: -lnL = 64.75 dispersal = 0.0471 extinction = 0.0668

At node a (origin of *Ruta*):

SPLIT	lnL	Rel.Prob
[Eu|Eu]	**-70.42**	0.003457
[Eu|Eu+WNAm]	-70.48	0.003276
[Eu|WNAm]	-70.63	0.002816
[An|An]	-70.64	0.002765
[Ae|Ae]	-70.79	0.002382
[Ib|Ib]	-70.94	0.002067
[Ap|Ap]	-71.09	0.001766
[Ar|Ar]	-71.15	0.001661

At node b (initial diversification of *Ruta*):

SPLIT	lnL	Rel.Prob
[Eu+Ae+An+Ba+Ib+Ap+Co+Sa+Ar|NAf]	**-68.21**	0.03143
[Eu+Ae+An+NAf+Ba+Ib+Ap+Co+Sa+Ar|NAf]	-68.26	0.03013

At node c (origin of the C-S lineage):

SPLIT	lnL	Rel.Prob
[Co|Co]	**-67.48**	0.06561
[Ap|Ap]	-68.16	0.03318

At node d (split between *R. corsica* and *R. lamarmorae*):

SPLIT	lnL	Rel.Prob
[Sa|Co]	**-65.43**	0.5114

At node e (origin of the Canarian endemics):

SPLIT	lnL	Rel.Prob
[NAf|NAf]	**-65.81**	0.3487

At node f (split of *R. oreojasme* from the rest):

SPLIT	lnL	Rel.Prob
[Go|Go]	**-66.56**	0.1637
[Te|Te]	-66.73	0.1384
[Lp|Lp]	-67.12	0.09418
[Go|Gc]	-67.37	0.07298
[Te|Gc]	-67.52	0.06308

At node g (split between *R. pinnata* and *R. microcarpa*):

SPLIT	lnL	Rel.Prob
[Te|Go]	**-65.93**	0.3075
[Lp|Go]	-66.11	0.2588

Chapter III

Phylogeny, morphology, and biogeography of *Haplophyllum* (Rutaceae), a species-rich genus of the Irano-Turanian floristic region

Gabriele Salvo[1], Sara Manafzadeh[1], Farrokh Ghahremaninejad[2], Komiljon Tojibaev[3], Louis Zeltner[4], Elena Conti[1]

[1]*Institute of Systematic Botany, University of Zürich, Zollikerstrasse 107, CH-8008 Zürich, Switzerland*.
[2]*Department of Biology, Tarbiat-Moaallem University, 49 Dr. Mofatteh Avenue, 15614 Tehran, Iran*.
[3]*Botanical Garden of the Academy of Sciences of Uzbekistan, Dj. Abidova str. 232, 100053 Tashkent, Uzbekistan*.
[4]*Laboratory of Evolutionary Botany, University of Neuchâtel, Emile Argand 11, CH-2007 Neuchâtel, Switzerland*.

Accepted for publication in:
Taxon

Abstract

Haplophyllum A. Juss. is one of the most species-rich, but poorly-known genera of Rutaceae (citrus family), reaching maximum species diversity in Turkey, Iran, and Central Asia. Many of its species exhibit a narrow geographic range (i.e., "narrow endemics"), which makes them particularly vulnerable to extinction. Despite its importance for the characterization of the Irano-Turanian floristic region, the evolution of species diversity in *Haplophyllum* has never been examined in a phylogenetic and biogeographic context. Consequently, we generated gene trees from DNA sequences of four regions of the chloroplast genome for 118 accessions, representing 66% of the species diversity of the genus. Additionally, *Haplophyllum* was examined morphologically. The phylogenetic analyses showed that several species of the genus do not form reciprocally-monophyletic groups. Optimization of morphological characters on the chloroplast DNA phylogeny indicated that most of the species, in particular those with a widespread geographic distribution, can only be diagnosed by combinations of homoplasious character states. Homoplasy notwithstanding, the main morphological characters traditionally used to classify the genus are consistent with the molecular phylogeny of *Haplophyllum*. Finally, the Mediterranean representatives of *Haplophyllum* were found to be embedded within a clade that includes primarily Irano-Turanian species, suggesting that multiple invasions of the Mediterranean basin from the east took place during the evolution of the genus.

Keywords

Phylogeny, morphology, biogeography, *Haplophyllum*, Irano-Turanian floristic region.

Introduction

Haplophyllum A. Juss. is one of the most species-rich, but poorly-known genera of Rutaceae. As currently circumscribed, it includes 68 species (Townsend, 1986; Navarro & al., 2004; Soltani & Khosravi, 2005) and reaches maximum species diversity in Turkey, Iran, and Central Asia (the latter region is bordered by the Caspian Sea in the west, China in the east, Iran and Afghanistan in the south, and Russia in the north). Many species of *Haplophyllum* exhibit a narrow geographic range (i.e., "narrow endemics"), a feature that makes them particularly vulnerable to extinction. Despite its importance for the characterization of the Irano-Turanian floristic region (Zohary, 1973; Takhtajan, 1986), the evolution of species diversity in *Haplophyllum* has never been examined in a phylogenetic and biogeographic context.

Haplophyllum is distributed from Morocco and Spain in the west to the Heilongjiang Province of China in the east. In the west it extends north to Romania and south to Somalia and the Hadhramaut area and in the east it extends north to the Lake Baikal region (Fig. 1; Townsend, 1986). Its range spans four different floristic regions: the Irano-Turanian, Mediterranean, Saharo-Arabian, and Sudano-Zambezian regions (Fig. 1; Takhtajan, 1986). The main centre of diversity of *Haplophyllum* is the Irano-Turanian region – in particular, Iran, Turkey, and Central Asia – which harbours 60% of the species diversity. Thirty species of *Haplophyllum* are present in Iran only, fourteen of which are endemic to the country (Joharchi, 2008). Fewer species occur in the other three floristic regions, most notably in the Mediterranean region, which contains 13% of the species diversity (Fig. 1).

A characteristic of many species of *Haplophyllum* is a highly restricted geographic distribution, sometimes consisting of a single mountain range (Townsend, 1986). For example, *H. telephioides* is found in a few mountains of central Anatolia; *H. viridulum* occurs in a small area of the Fars province of Iran; and *H. eugenii-korovinii* is restricted to the Karatau mountains of Kazakhstan, where it is very rare (Townsend, 1986). Overall, 54% of the species have a relatively narrow range as compared to the most widespread species, which constitute 18% of the total; the remaining species exhibit an intermediate distribution. Additionally, several endemic species of *Haplophyllum* occur in small, disjunct populations across their narrow range (Gabriele Salvo, Sara Manafzadeh, pers. obs.). These factors make many species and populations of the genus potentially in danger of extinction, a fact that has been recognized with the inclusion of nine species in the Red Data Book of Iran (Jalili & Jamzad, 1999). Conversely, some species of *Haplophyllum* have a very widespread distribution. For example, *H. tuberculatum* stretches from Morocco to western Pakistan,

broadly spanning the distribution of the entire genus; *H. buxbaumii* is found from Morocco to western Iran, including many islands in the eastern Mediterranean Sea (Townsend, 1986).

Haplophyllum species are perennial herbs, sometimes low shrubs, which grow mainly on sandy, stony, or rocky hill slopes in arid areas (Townsend, 1986). Morphologically, they can be broadly characterized by the presence of cymose and bracteate inflorescences with five sepals and creamy-white to bright-yellow petals, ten stamens with free filaments expanded below and pubescent on the inner surface, three to five connate carpels, and five-lobed capsules (Townsend, 1986).

Haplophyllum has been studied from a morphological (Jussieu, 1825; Spach, 1849; Boissier, 1867; Engler, 1896; Vvdensky, 1949; Townsend, 1986) and phytochemical (e.g., Mester & Vicol, 1971; Pascual-Villalobos & Robledo, 1999; Shaiq & al., 2001; Nazrullaev & al., 2002; Prieto & al., 2002) point of view. The most comprehensive morphological analysis of the genus was published by Townsend (1986), who also proposed a classification and a tentative scheme of species relationships. Phytochemically, Rutaceae as a whole are notable for their vast array of secondary chemical compounds (e.g., alkaloids, lignanes, glycosides, flavonoids; Price, 1963). Mester & Vicol (1971) performed a thorough phytochemical analysis of *Haplophyllum* by focusing on the distribution of different classes of alkaloids. However, on the basis of these two sources of data – morphology and phytochemistry – both the generic status of the genus and its subdivision into different sections have been questioned.

In the most comprehensive classification of Rutaceae, based mainly on morphological characters, Engler (1896, 1931) treated *Haplophyllum* as a subgenus of *Ruta* L. This view was dismissed by subsequent systematic works, which emphasized the distinctiveness of the former taxon with respect to both morphological and phytochemical features. Townsend (1986) listed a series of morphological traits that can differentiate *Haplophyllum* from closely-related genera (e.g., pollen structure, seed shape, petal margins). Mester & Vicol (1971) discovered the presence of secondary metabolites in *Haplophyllum*, such as the alkaloids robustine, haplopine, and skimmianine, which are not found in any other genus of Rutaceae.

Several systematists have attempted to subdivide *Haplophyllum* into different sections by means of morphological characters (Spach, 1849; Boissier, 1867; Engler, 1896; Vvdensky, 1949; Townsend, 1986). Of these, only the last two authors adopted explicit criteria, rather than generic statements, to support their classifications. Vvedensky (1949) divided the genus into four sections according to carpel number, fruit opening, and ovule number (Table 1). The first two features, together with petal colour, plant architecture, ovary shape, and stamen

form, were used by Townsend (1986) to divide the genus into three sections (Table 1). In his assessment of the taxonomic value of different morphological characters, Townsend (1986) noted that the ovary and stamens provide the most useful characters to infer species relationships within *Haplophyllum*. However, both classifications have been criticized, because they lack morphological traits that are consistent across all species of the proposed sections (Mester & Vicol, 1971).

The single phylogenetic study of *Haplophyllum* available so far included only six of the 68 species and focused on the Iberian representatives of the genus (Navarro & al., 2004). More recently, Salvo & al. (2008) performed a phylogenetic analysis of tribe Ruteae, which includes *Haplophyllum* and closely-related genera, based on chloroplast (cp) DNA sequences. This study, comprising a limited sample of 22 species of *Haplophyllum*, corroborated the monophyly of the genus, but did not address species relationships within it.

From a biogeographic point of view, *Haplophyllum* was used to characterize the Irano-Turanian region (Zohary, 1973, in relation to his Western Irano-Turanian subregion; Takhtajan, 1986, in relation to his Western Asiatic subregion), because many of its species are restricted to this geographic area. For similar reasons, Grubov (1959) mentioned *Haplophyllum* in the characterization of Central Asia. In its most common delimitation, the Irano-Turanian region extends from central and eastern Anatolia to the Tien Shan and Altai mountain ranges, reaching the Gobi desert, and includes parts of the Sinai peninsula, Lebanon, Jordan, Israel, and Palestine, most of Syria and Iran, northern Iraq, north-eastern Afghanistan, parts of northern Pakistan and northern India, and Central Asia (Fig. 1; Takhtajan, 1986; Davis & al., 1994). Based on either floristic similarities or phylogenetic evidence, some authors suggested that the Irano-Turanian region served as a key source for the colonization of neighbouring areas, most notably the Mediterranean region (Zohary, 1973; Quézel, 1978, 1985, 1995; Ribera & Blasco-Zumeta, 1998; Thompson, 2005; Mansion & al., 2008, 2009), while others argued more generally that the present arid floras of Eurasia, the Mediterranean region, North Africa, and even South Africa originated from Central Asia (Bobrov, 1965, 1966; Pyankov & al., 2002).

While an excellent taxonomic monograph on *Haplophyllum* is available (Townsend, 1986) and a detailed knowledge of the evolution of species diversity in this taxon could yield useful insights into the biogeographic role of the Irano-Turanian region, the genus has never been comprehensively examined from a phylogenetic/biogeographic point of view. In order to fill this gap of knowledge, we generated sequence data for 66% of the species diversity of *Haplophyllum* and addressed the following questions: 1) Are the different species of *Haplophyllum* monophyletic? 2) Does our inferred cpDNA phylogeny support Vvedensky's

(1949) or Townsend's (1986) classifications? 3) Do species from the same floristic region form monophyletic groups? 4) Did the Irano-Turanian region serve as a source for the colonization of the Mediterranean region? 5) What are the phylogenetic relationships between the narrow endemics and the geographically-widespread species?

Material and Methods

Taxon sampling

Forty five out of 68 species of *Haplophyllum* were sampled, including species with a very narrow distribution occurring in remote areas. For geographically-widespread taxa, and/or taxa that are difficult to diagnose morphologically, multiple accessions per species (two to eleven) were sampled. All three sections of Townsend (1986) were sampled. Five outgroup taxa were selected according to previous phylogenetic results: *Cneoridium dumosum*, *Aegle marmelos*, *Citrus reticulata*, *Poncirus trifoliata*, and *Glycosmis citrifolia* (Salvo & al., 2008). The final matrix contained 118 accessions. Included material, voucher information, sources, and GenBank/EBI accession numbers are listed in Appendix 1.

Character sampling

To allow for inclusion of the new molecular data in a global dataset of Rutaceae, the following cpDNA markers were chosen: the *mat*K gene, the *trn*K gene, the *rpl*16 intron, and the *trn*L-*trn*F intergenic spacer. These markers enabled us to produce unequivocal alignments and provided sufficient resolution at our level of investigation.

DNA extraction, amplification and sequencing

Prior to DNA extraction, silica-dried leaf material (15-20 mg) was ground using glass beads and a MM 3000 shaker (Retsch GmbH, Haan, Germany). Total genomic DNA was extracted using DNeasy Plant Mini Kits from QIAGEN AG (Basel, Switzerland), following the manufacturer's instructions. The *mat*K and *trn*K cpDNA coding regions were amplified using primers 1F and 1R (Sang & al., 1997). The *rpl*16 intron was amplified using primers F71 and R1516 (Baum & al., 1998). The *trn*L-*trn*F spacer was amplified with primers c and f (Taberlet & al., 1991). All PCR reactions were 20 µl in volume. Each reaction included 9.2 µl of ddH_2O, 2 µl of Taq-Buffer [10x, 15mM $MgCl_2$], 1.6 µl of $MgCl_2$ [25 mM], 3.2 µl of dNTP [1.25 mM], 0.2 µl of Taq-Polymerase [5U/µl], 1 µl of BSA, 0.4 µl of each primer (forward and reverse), and 2 µl of DNA template. Amplification of the *mat*K region consisted of 2 min at 94°C followed by 30 cycles of: 1.5 min denaturation (94°C), 2 min annealing (53°C), and 3 min extension (72°C). After the last cycle, the temperature was kept at 72°C for the last 15 min of extension and then lowered to 4°C. Amplification of both the *rpl*16 and *trn*L-*trn*F regions consisted of 2 min at 94°C followed by 35 cycles of: 0.5 min denaturation (94°C), 1

min annealing (52°C), and 1.75 min extension (72°C). After the last cycle the temperature was kept at 72°C for 10 min of extension and then lowered to 4°C. All PCR and cycle sequencing reactions were run on a TGradient thermocycler (Biometra, Goettingen, Germany). In order to detect amplified DNA target regions and possible contamination, PCR products were separated on 1% agarose gels, stained with ethidium bromide, and viewed under UV light. Successfully amplified products were purified with the GFX PCR DNA and Gel Band purification Kit (Bioscience Amersham, Otelfingen, Switzerland), following the manufacturer's recommendations.

Cycle sequencing reactions were carried out using the BigDye™ Terminator Mix (Applied Biosystems, Inc., Foster City, California) and the same primers as above. The sequencing protocol consisted of 24 cycles of 10 sec denaturation (96°C), 5 sec annealing (50°C), and 4 min elongation (60°C). Products were run on an ABI 3100 Genetic Analyzer (Applied Biosystems, Inc., Foster City, California) according to the manufacturer's protocol. For each region both strands were sequenced.

Phylogenetic analyses

Sequences were edited and assembled using Sequencher 4.2™ software (Gene Codes Corp., Ann Arbor, Michigan, USA). Base positions were individually double-checked for agreement between the complementary strands. All sequences were visually aligned in MacClade 4.06 (Maddison & Maddison, 2000) using the similarity criterion (e.g., Simmons, 2004). Regions of ambiguous alignment were excluded from the analysis (Kelchner, 2000). Gap positions were treated as missing data, unequivocally aligned gaps being coded as presence/absence of characters with the software GapCoder (Young & Healy, 2003) and then added as binary characters to the data matrix.

Four data partitions were defined, corresponding to the four loci of the chloroplast genome examined in this study. The individual partitions were initially analyzed separately to establish whether there were any well-supported, incongruent clades among the respective trees. Since no such incongruence was detected, the sequences of the four loci were combined in a single dataset. The combined matrix was analyzed using both Maximum Parsimony (MP) and Bayesian MCMC Inference (BI; Yang & Rannala, 1997). Parsimony analyses were conducted using PAUP*4.0b10 (Swofford, 2001). All changes were treated as unordered and equally weighted (Fitch, 1971). Tree search was performed using the following protocol: (i) a heuristic search was carried out with 1000 replicates of random taxon addition sequence and 10 trees held at each step, and tree bisection-reconnection branch swapping (TBR) on best trees only, with no more than 100 trees saved per replicate; (ii) the best trees found in (i) were then used as starting trees for a second heuristic search using TBR branch swapping until all

swapping options were explored, and saving multiple trees (MULTREES option in effect). The STEEPEST DESCENT option was used in both (i) and (ii). Relative support for each node obtained by MP was assessed using bootstrap re-sampling (Felsenstein, 1985). The following protocol was employed: heuristic search, 1000 bootstrap replicates, ten random addition sequence replicates with three trees held at each step, TBR swapping with STEEPEST DESCENT and saving no more than 50 trees per replicate.

Bayesian inference of phylogeny was performed with MRBAYES v3.1.2 (Ronquist & Huelsenbeck, 2003). First, the model of evolution most suitable for each individual cpDNA region was determined with the Akaike Information Criterion (AIC; Akaike, 1974) in ModelTest 3.06 (Posada & Crandall, 1998). Subsequently, five partitions corresponding to the four loci (only nucleotide characters) and the coded gap characters were specified, and the commands "lset NST=6, RATES=gamma" and "lset coding=variable" were entered in MRBAYES v3.1.2 for the former and the latter, respectively. Six independent runs with four Monte Carlo Markov chains (one cold and three incrementally heated; TEMP = 0.1) run for 5 x 10^6 generations each, with trees sampled every 1000^{th} generation, were performed. The first 1 to 2.5 x 10^6 sampled trees of each run were discarded as "burn-in", after checking for stability on the log-likelihood curves using the software Tracer v1.4 (Rambaut & Drummond, 2007) and after visual inspection of the split (clade) frequencies using the software AWTY (Wilgenbusch & al., 2004; http://ceb.csit.fsu.edu/awty). The remaining 22000 trees were used to build a 50% majority rule consensus tree.

Morphological data

A matrix was constructed for 27 discrete morphological characters scored using herbarium material (Z, G, P, MA, W, TBI, LE, TAK, FAR, TARI) and Townsend's (1986) monograph, for the same 45 species of *Haplophyllum* used in the phylogenetic analyses and for its sister group, *Cneoridium dumosum* (Appendix 2, 3). These characters represent vegetative (characters 1-10), inflorescence (including both stamen and pistil features; 11-26), and fruit (27) morphology (Appendix 2). When possible, morphological characters were assessed for several specimens of each species. All characters were treated as unordered; 23 characters were binary and 6 were multistate (Appendix 2). Autapomorphies were not included in the matrix. Missing data and polymorphic character states represented 3.2% and 1.1% of the data matrix entries, respectively (Appendix 3).

Morphological analyses

Initially, the matrix was analysed using cladistic methods in PAUP*4.0b10; however, the resulting tree was poorly resolved and weakly supported, even after tree searching was performed using successive weighting (Farris, 1969; results not shown). This is a known

problem of reconstructing phylogenies using morphological data only (Scotland & al., 2003). Since the morphological matrix consisted of categorical data, a multiple correspondence analysis (MCA; Benzecri, 1992; Venables & Ripley, 2002) was carried out using the statistical software-package SPSS for Windows Rel. 11.0.1 (SPSS Inc., 2001), in order to visualize the joint properties of the 27 morphological variables in two dimensions.

Character mapping analyses

To assess the fit of each morphological character onto the inferred molecular phylogeny, all morphological characters were mapped onto a subset of the post-"burn-in" Bayesian trees. The subset was created by sampling a tree every 100 trees from the original set of trees, yielding a total of 220 trees. The 50% majority rule consensus of these trees was identical to the one from the original set of trees, indicating that our subset was representative of the original set of trees. Four taxa belonging to the outgroup were pruned from the 220 trees, leaving only the sister group of *Haplophyllum*, *Cneoridium dumosum*. The fit of each character onto a tree was assessed using the rescaled consistency index (RC; Farris, 1989). This index has been shown to be superior to both the consistency and retention indexes in assessing fit of characters onto a phylogeny (Kitching & al., 1998). The character mapping analyses were implemented in PAUP*4.0b10 and MESQUITE ver. 2.7.1 (Maddison & Maddison, 2008) using parsimony as the optimization procedure and treating character state transitions as unordered.

Results

Phylogenetic analyses

The combined molecular matrix consisted of 3849 characters, of which 561 were parsimony-informative. The MP analysis yielded 9400 most parsimonious trees of 1485 steps, with a consistency index (CI) of 0.66 and a retention index (RI) of 0.86. The AIC, as implemented in ModelTest, selected the following models of evolution: GTR+G for the *mat*K region, TVM+G for both *rpl*16 and *trn*L-*trn*F, and TIM+G for *trn*K. The 50% majority-rule consensus tree obtained from the Bayesian analysis is shown in Fig. 2a. This tree is slightly more resolved than the strict consensus tree found from the MP analysis of the same matrix. Branch support values, in terms of both bootstrap percentages (BP) and posterior probabilities (PP), were generally lower along the backbone of the tree and higher towards the tips. Two main strongly-supported (i.e., BP > 69 and PP > 0.94; Hillis & Bull, 1993; Zander, 2004) clades can be identified: clade A, including only Irano-Turanian species, such as the characteristic *H. acutifolium* and *H. robustum*, and clade B, containing species from different floristic regions, including Mediterranean representatives and also the widespread species *H.*

buxbaumii and *H. tuberculatum* (Fig. 2a). Many species represented by multiple accessions were either poorly resolved or non-monophyletic, but only a few cases of non-monophyly were strongly supported. Neither the Irano-Turanian and Mediterranean representatives nor the species occurring in more than one floristic region formed monophyletic groups (Fig. 2a).

Morphological analyses

The results of the MCA are displayed in Fig. 2b. The first and second dimensions explained 23% and 16% of the total variance, respectively. The characters that contributed the most to the first and second dimensions were characters 6 (0.603), 13 (0.557), 23 (0.514), and characters 13 (0.663), 15 (0.638), 22 (0.630), 18 (0.519), respectively (Appendix 2).

Character mapping analyses

The results of the character mapping analyses are summarized in Fig. 3 (see also Appendix 2). A lot of variation in mean RC values was found across characters, with "stem branching" (character 5, RC = 0.024) and "number of carpels" (character 21, RC = 0.257) receiving the lowest and highest value, respectively. In the vegetative-morphology category, the characters that showed the best fit onto the tree were "sterile axillary shoots" (character 3, RC = 0.136) and "number of stems" (character 1, RC = 0.128). In the inflorescence-morphology category, the features that received the highest RC values were: "number of carpels" (character 21, RC = 0.257), "number of ovules" (character 24, RC = 0.193), "indumentum of filament" (character 18, RC = 0.130), and "dark dorsal vitta/tinge on petal" (character 14, RC = 0.116). Overall, the characters that Vvedensky (1949) and Townsend (1986) valued the most in their classification of *Haplophyllum* (Table 1), namely "number of carpels", "number of ovules", and "carpel dehiscence" (character 27, RC = 0.238), showed the best fit onto the molecular tree (Fig. 3). Details of the mapping of these three characters onto the molecular phylogeny are shown in Fig. 4.

Discussion

The main goal of our study was to provide an initial estimate of the evolutionary history and taxonomic relationships of *Haplophyllum* derived from cpDNA sequences and morphology. To define species boundaries we used the conceptual framework of the phylogenetic species concept *sensu* Baum (1992). According to this concept: "taxa (including species) are viewed as monophyletic or exclusive groups of organisms" (Baum & Donoghue, 1995, p. 569). In this view, species are defined within the historical dimension provided by the phylogeny at hand (i.e., diachronistically) and in order to count as natural entities must be monophyletic (Rieppel, 2010). While we are aware of the debate on multiple species concepts (e.g., Hennig, 1966; Cracraft, 1989; Baum & Donoghue, 1995; Freudenstein, 1998; Rieppel,

2010), an exhaustive discussion of the pros and cons of each concept is beyond the goal of the present paper.

Systematics

The phylogenetic analyses indicated that several taxonomic species of *Haplophyllum* were non-monophyletic. Even though most of the inferred cases of non-monophyly were weakly supported, possibly resulting at least in part from inadequate phylogenetic signal (Syring & al., 2007), a few instances of species-level paraphyly and polyphyly were strongly supported (Fig. 2a). These raise concerns about species circumscription within *Haplophyllum* and potential discrepancies between gene trees and species trees. Commonly-cited, causative factors responsible for species non-monophyly and gene tree versus species tree incongruence are: imperfect taxonomy (e.g., Goodwillie & Stiller, 2001), introgressive hybridization (e.g., Shaw & Small, 2005), incomplete lineage sorting (e.g., Bouillé & Bousquet, 2005), unrecognized amplification of paralogous loci (e.g., Alvarez & al., 2005), and recombination among divergent alleles (e.g., Schierup & Hein, 2000). Although the phylogeny inferred in the present study is based on cpDNA markers only, making it difficult to assess the relative contribution of the above-mentioned factors to our study-group, our phylogenetic results, coupled with evidence from morphology, distribution, and ecology, represent a useful first step towards addressing the issue of species circumscription and identity in *Haplophyllum*.

H. tuberculatum, *H. buxbaumii*, *H. virgatum*, and *H. blanchei* were inferred to be non-monophyletic, although branch support in the clade that includes their accessions was generally low (Fig. 2a). Moreover, the level of intra-specific polymorphism within these taxonomic species was similar to the level of inter-specific divergence between them (Table 2). For example, the average absolute number of nucleotide substitutions between the accessions of *H. buxbaumii*, and between these and the accessions of *H. tuberculatum*, *H. virgatum*, and *H. blanchei*, was 13.5 and 14.5, respectively. Such levels of intra-specific polymorphism are similar to those found in other studies at the genus level that have sequenced multiple infra-specific accessions (e.g., Särkinen *et al*., in press; Widmer & Baltisberger, 1999). *H. tuberculatum*, *H. buxbaumii*, and *H. blanchei* are morphologically similar (Fig. 2b). *H. tuberculatum* and *H. buxbaumii* exhibited the same character states for 17 out of the 27 scored morphological characters (Appendix 3). They are the species with the most widespread distribution within *Haplophyllum* and with the highest level of intra-specific morphological variability, which led Townsend (1986) to recognize two "morphs" within *H. tuberculatum* and two subspecies in *H. buxbaumii*. As a matter of fact, Townsend (1966a, p. 99) stated that the circumscription of *H. tuberculatum* is "the most difficult problem to be solved in the genus". *H. blanchei* is difficult to separate from *H. tuberculatum* on the basis of

morphology, the chief separating characters being the bright-magenta-coloured flowers and distinctly-fused filaments of *H. blanchei* (Townsend, 1986). Furthermore, the geographic ranges of these two species overlap. The taxonomic status of *H. virgatum* is unclear (Townsend, 1986). From a morphological standpoint, this species is difficult to separate from *H. canaliculatum*. In fact, in Flora Iranica Townsend (1966b) reduced these two species to synonymy. In the morphological matrix, although *H. virgatum* shared the same character state with *H. canaliculatum* with respect to three vegetative characters, it possessed the same character state as *H. tuberculatum*, *H. buxbaumii*, and *H. blanchei* for eight characters (Appendix 3).

Within this problematic clade, including *H. tuberculatum*, *H. buxbaumii*, *H. virgatum*, *H. blanchei*, *H. laristanicum*, and *H. dasyginum* (Fig. 2a), the only strongly-supported relationships were found for a sister relationship between one accession of *H. buxbaumii* and *H. laristanicum* (74 BP, 1.0 PP), and for a clade including *H. blanchei* and one accession of *H. tuberculatum* and *H. buxbaumii* each (72 BP, 1.0 PP). *H. laristanicum* is a rare narrow endemic restricted to a small part of southern Iran. Its range is included within the range of *H. tuberculatum*, but it is geographically separated from the range of *H. buxbaumii*. Morphologically, *H. laristanicum* bears a resemblance to some forms of *H. tuberculatum*, which have the filaments almost as fused (Townsend, 1986). In the morphological matrix, *H. laristanicum* shared an identical character state with *H. buxbaumii* and *H. tuberculatum* for four characters, respectively (Appendix 3).

In conclusion, due to the interdigitation of accessions from different species, similar levels of intra-specific polymorphism and inter-specific divergence, high levels of intra-specific morphological variability, morphological character conflict, and overlapping geographic ranges, species circumscription and phylogenetic relatedness within this problematic clade remains unclear. Taxonomic "lumping" may partially be responsible for the inferred phylogenetic interdigitation of the accessions of *H. tuberculatum* and *H. buxbaumii*, due to the possible presence of unrecognized, "cryptic" species within the range of both species. A much more thorough infra-specific sampling, including several populations per species and individual samples covering the entire geographic distribution of each species, in order to fully represent the extent of morphological variation, is needed to disentangle species relationships within this "species complex".

The main finding of the MCA was a sharp morphological separation between *H. furfuraceum* and *H. erythraeum*, and the remaining species of the genus (Fig. 2b). These two species share the presence of a characteristic farinose indumentum on the sepals, petals, and ovary (Appendix 2). However, in the inferred phylogeny, *H. furfuraceum* and *H. erythraeum*

did not form a monophyletic group, but were intermingled with *H. affine* and *H. glaberrimum* (Fig. 2a). In Townsend's (1986) tentative scheme of species relationships *H. furfuraceum*, *H. erythraeum*, and *H. affine* are clustered together, whereas *H. glaberrimum* is placed in a very distant position from these three species.

The lack of monophyly in *H. bungei*, *H. alberti-regelii*, and *H. versicolor* is strongly supported: one accession of *H. bungei* was inferred to be sister to two accessions of *H. versicolor* with 88 BP and 1.0 PP, whereas the remaining two accessions of *H. bungei* were part of a strongly-supported clade comprising representatives of *H. alberti-regelii* and *H. dubium* (83 BP, 1.0 PP; Fig. 2a). From a morphological point of view, *H. bungei* and *H. alberti-regelii* are more similar to each other than they are to *H. versicolor* (Fig. 2b). Therefore, the accession of *H. bungei* inferred to be sister to *H. versicolor* seems to be in a spurious phylogenetic position.

Another case of strongly-supported species non-monophyly is represented by *H. thesioides*. One accession of this species was inferred to be sister to *H. telephioides* with strong support (100 BP, 1.0 PP), whereas for the other accession a sister relationship with *H. suaveolens* was strongly supported (87 BP, 1.0 PP; Fig. 2a). Morphologically and ecologically, *H. thesioides* is different from *H. telephioides* but similar to *H. suaveolens* (Fig. 2b). Townsend (1986; p. 297) stated: "This species [*H. thesioides*] and *H. suaveolens* have been much confused in herbaria." *H. telephioides* exhibits a characteristic plant architecture and petal coloration, not encountered in the other two species, and is confined to rocky, limestone slopes, whereas *H. thesioides* and *H. suaveolens* have much broader habitat preferences. Moreover, the accession of *H. thesioides* inferred to be sister to *H. suaveolens* was collected from an area where the latter species occurs, but where *H. telephioides* is absent. Hence, it is possible that this accession represents a case of introgressive hybridization between *H. thesioides* and *H. suaveolens*. Such a process has already been proposed for the origin of *H. ptilostylum* (Townsend, 1966a).

The phylogenetic results also identified several instances of strongly-supported species monophyly: *H. dauricum* (100 BP, 1.0 PP), *H. griffithianum* (94 BP, 1.0 PP), *H. robustum* (73 BP, 1.0 PP), *H. latifolium* (92 BP, 1.0 PP), *H. pilosum* (100 BP, 1.0 PP), *H. telephioides* (98 BP, 1.0 PP), *H. coronatum* (100 BP, 1.0 PP), and *H. albanicum* (93 BP, 1.0 PP; Fig. 2a). Some of these species, especially the narrow endemics (see below), can be diagnosed by a set of morphological features, which, however, are not exclusive to them. For example, *H. dauricum* is a small, suffrutescent plant and has a distinctive plant architecture, with numerous, slender stems arising from a stout, woody base; features that are also encountered in *H. bucharicum*. Together with *H. gilesii*, this species is the only one with a 3-

locular (rarely 2- or 4-5-locular) ovary (Table 1). *H. robustum* is easy to recognize in the field, due to its broad lanceolate leaves and stout, erect stems, usually un-branched below the inflorescence, reaching up to 80 cm in height (Gabriele Salvo, Sara Manafzadeh, pers. obs.). Such characteristics are also encountered in *H. latifolium* and *H. popovii*. The typical lanate indumentum of *H. pilosum* occurs also in *H. villosum*, *H. telephioides*, *H. suaveolens*, and *H. coronatum*, although to a lesser extent. The prominent, apical appendages found in the ovary of *H. albanicum* individuals are also present in *H. coronatum*, *H. broussonetianum*, *H. pilosum*, *H. telephioides*, *H. suaveolens*, *H. linifolium*, *H. balcanicum*, *H. armenum*, and *H. patavinum*, and others, although the shape of this feature varies slightly among species (Townsend, 1986). In essence, only by means of combinations of homoplasious, morphological character states (i.e., changing more than once across the *Haplophyllum* phylogeny) are we able to diagnose the different species of the genus. In fact, the character mapping analyses detected high levels of homoplasy across most of the scored morphological characters (Fig. 3). A similar situation has been found in other taxonomically-complex plant groups (e.g., Moylan & al., 2004; Norup & al., 2006).

Haplophyllum species with a widespread distribution are often more difficult to diagnose morphologically as compared to narrow endemics. For example, the degree of fusion of the filaments, which is a very important character for the classification of the genus, is variable only in the broadly-distributed *H. tuberculatum*, which includes individuals with filaments that are either free or joined at the base (Townsend, 1986). Likewise, in *H. buxbaumii* the form of the apex of the ovary, which is another crucial taxonomic character, is variable, with individuals either lacking or possessing an apical appendage on the ovary (Townsend, 1986). On the contrary, species with a narrow range, such as *H. telephioides* or *H. bucharicum*, can be easily diagnosed by clear morphological features, even though sometimes these are also present in a few other species. A dark-green line along the dorsal side of the petals, for example, is very prominent in *H. telephioides*, although not restricted to it (Townsend, 1986). Similarly, a woody, frequently gnarled base of the stem is very distinct in *H. bucharicum* (Townsend, 1986).

Such extensive morphological polymorphism in some broadly-ranging species groups has long been observed by taxonomists and has posed several problems for the delimitation of species boundaries (Mayr, 1942; Wilson & Brown, 1953). This common observation represents an interesting link between biogeography and systematics. It is likely that narrow endemics are more specialized in their ecological requirements, as compared to species with a widespread distribution. This specialization may consist in the acquisition of unique

morphological features (i.e., apomorphic character states), which are the result of adaptation to local environmental conditions and make the narrow endemics "diagnosable".

Neither Vvedensky's (1949) nor Townsend's (1986) classifications of *Haplophyllum* were supported by our phylogenetic findings (Table 1, Fig. 2a). *H. acutifolium* and *H. latifolium*, placed by both systematists in the same section, were not inferred to be sister to one another. The former was found to be sister to *H. robustum*, although with low support (< 50 BP, 0.52 PP); the latter exhibited a strongly-supported sister relationship to *H. popovii* (93 BP, 1.0 PP). Additionally, Townsend's (1986) section *Haplophyllum* and Vvedensky's (1949) sections *Polyoon* and *Oligoon* did not form monophyletic groups. Unfortunately, the validity of section *Peganoides sensu* Townsend (1986) could not be ascertained, since we were unable to obtain samples of *H. gilesii*, a species endemic to the Kashmir region.

The main morphological characters used by both systematists to divide the genus into sections – namely, "number of carpels", "capsule dehiscence", and "number of ovules" – exhibited the lowest levels of homoplasy when optimized on the inferred phylogeny (Fig. 3). For example, *H. dauricum* and *Cneoridium dumosum* are the only sampled taxa that do not have five carpels, whereas all the others do (Fig. 4, Appendix 2, 3). Similarly, *H. acutifolium*, *H. latifolium*, and *Cneoridium dumosum* are the only taxa with indehiscent capsules (Fig. 4, Appendix 2, 3). "Number of ovules", which was used for classificatory purposes by Vvedensky (1949) only, shows a more complex pattern; however, in this case too, the least common character state (more than four ovules) is only found in *H. pilosum* and *H. broussonetianum*, and the next, less common state (four ovules) exhibits great phylogenetic structure (Fig. 4, Appendix 2, 3). Such unbalanced distribution of characters states, with most of the taxa represented by one state and only a few taxa by the other state(s), means that opportunities for state transitions are few and hence levels of homoplasy low (Sanderson & Donoghue, 1989). Overall these findings emphasize the important taxonomic value of the three mentioned characters within *Haplophyllum*. More generally, morphological features of the inflorescence and fruit provide the most useful taxonomic characters to infer species relationships within the genus (Fig. 3). This fact was noted by Townsend (1986, p. 3) who stated: "The ovary furnishes some of the most useful characters in classifying the genus".

Biogeographic patterns

The phylogenetic results indicated that species from the same floristic region do not form monophyletic groups (Fig. 2a). Even though the Mediterranean representatives of the genus do not cluster together, they are embedded within a clade that includes primarily Irano-Turanian species and species that occur in more than one floristic region (Fig. 2a), suggesting that multiple invasions of the Mediterranean region from the east took place during the

evolution of the genus. A pattern of migration from western Asia into the Mediterranean basin has been inferred for the origin of *Arum* and *Biarum*, two genera of Araceae restricted primarily to the Mediterranean region (Mansion & al., 2008). Similarly, an Anatolian origin has been inferred for *Anchusa*, *Borago*, and *Echium*, three genera of Boraginaceae that comprise members endemic to the western Mediterranean region (Mansion & al., 2009).

Both the geographically-widespread species and the narrow endemics were found to be intermingled across the phylogeny (Fig. 2a). A few terminal clades containing a widespread species and a narrow endemic were inferred: *H. dauricum*/*H. dshungaricum* (although weakly supported); *H. boissierianum*/*H. albanicum* (strongly supported); *H. myrtifolium*/*H. cappadocicum*, *H. thesioides*/*H. telephioides*, and *H. buxbaumii*/*H. laristanicum* (although the widespread species of these three clades were non-monophyletic; Fig. 2a). An expanded taxon sampling within such clades, especially with respect to the narrow endemics, will enable us to verify whether the narrow endemics and the widespread taxa form reciprocally-monophyletic sister pairs or whether the narrow endemics are nested within paraphyletic widespread taxa. These different phylogenetic patterns have implications on the origin of the narrow endemics and the geography of speciation (e.g., Bush, 1975; Lynch, 1989; but see Losos & Glor, 2003). The former pattern is compatible with an allopatric mode of speciation if the sister species display little or no overlap in their geographic ranges or with a sympatric mode of speciation if the geographic ranges of the sister species overlap (Barraclough & Vogler, 2000). The latter pattern would point to a peripatric mode of speciation if the narrow endemic species is geographically isolated from the widespread species (e.g., Harrison, 1991).

For example, the widespread *H. boissierianum* and the narrow endemic *H. albanicum* exhibit a strongly-supported sister relationship (Fig. 2a) and an overlapping geographic range (the range of the latter species is contained within the range of the former one), suggesting that the two species originated *via* a sympatric mode of speciation. This view is supported by their different ecological preferences: *H. boissierianum* occurs on rocky and stony places, on hill slopes, along roads, in open *Pinus* woodlands, and on limestone or serpentine soil, whereas *H. albanicum* is restricted to rocky and stony habitats with limestone soil (Townsend, 1986).

Conclusion

The present study represents a first step towards disentangling species relationships in a taxonomically-complex and biogeographically-important genus by means of phylogenetic and morphological analyses. The phylogenetic analyses identified both cases of strongly-

supported species monophyly and instances of species non-monophyly. The morphological assessment showed that the different species of the genus, especially those with a widespread distribution, cannot be readily diagnosed by sets of unique character states. Character mapping analyses indicated that the main morphological characters traditionally used to classify the genus are consistent with the molecular phylogeny of *Haplophyllum*. Preliminary biogeographic patterns suggest that the Mediterranean representatives of *Haplophyllum* arrived from the east multiple times.

The inferred phylogenetic framework will lay the foundations for future studies that will focus on selected, problematic clades (e.g., the *H. tuberculatum* / *H. buxbaumii* clade). Within such clades it will be possible to expand the current infra-specific sampling and perform more detailed molecular (examining haplotype variation, for example) and morphological analyses, which are necessary to determine species boundaries and test the validity of the different species of *Haplophyllum*. Additionally, sampling the nuclear genome will be a requisite in order to understand the biological processes underlying species non-monophyly in *Haplophyllum*.

Acknowledgments

The authors would like to thank the following people: Guilhem Mansion; Federico Selvi; Oyumaa Jamsran; Shahin; Manana Khutsishvili; Malvina Davlianidze; the staff of the LE, P, TAK, G (in particular, Nicolas Fumeaux), and W (in particular, Ernst Vitek) herbaria; Simon Ho; Frédéric Médail; Barbara Keller; Jurriaan de Vos; Peter Endress; Rolf Rutishauser; Alexandre Antonelli; Aelys Humphreys; Colin Hughes; Shuqing Xu; Daniele Cicuzza; Reto Nyffeler; Elena Benetti; Barbara Seit; Victoria Smith; Mary Endress and two anonymous reviewers. This study was funded by the Swiss National Science Foundation grant no. 3100A0-105495 to E.C., and the Institute of Systematic Botany and International Student Office of the University of Zürich. Additional financial support from the Georges-und-Antoine-Claraz-Schenkung, the Kommission für Reisestipendien der Schweizerischen Akademie der Naturwissenschaften (SANW), the Equal Opportunity Office of the University of Zürich, and the Synthesys Programme of the EU is gratefully acknowledged.

References

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Trans. Automat. Control* 19: 716–723.

Álvarez, I., Cronn, R., & Wendel, J.F. 2005. Phylogeny of the New World diploid cottons (*Gossypium* L. Malvaceae) based on sequences of three low-copy nuclear genes. *Plant Syst. Evol.* 252: 199–214.

Barraclough, T.G., & Vogler, A.P. 2000. Detecting the geographical pattern of speciation

from species-level phylogenies. *Am. Nat.* 155: 419–434.
Baum, D.A. 1992. Phylogenetic species concepts. *Trends Ecol. Evol.* 7: 1–2.
Baum, D.A., & Donoghue, M.J. 1995. Choosing among alternative "phylogenetic" species concepts. *Syst. Bot.* 20: 560–573.
Baum, D.A., Small, R.L., & Wendel, J.F. 1998. Biogeography and floral evolution of baobabs (*Adansonia*, Bombacaceae) as inferred from multiple data sets. *Syst. Biol.* 47: 181–207.
Benzecri, J.-P. 1992. *Correspondence Analysis Handbook*. New York: Marcel Dekker.
Bobrov, E.G. 1965. On the origin of the Old World deserts, as illustrated by the genus *Nitraria* L. *Bot. Zhurn.* 50: 1053–1067.
Bobrov, E.G. 1966. A review of the genus *Reaumuria* L. in connection with the problem of the origin of the Afro-Asiatic desert flora. *Bot. Zhurn.* 51: 1057–1072.
Boissier, E. 1867. Rutaceae. Pp. 920–943 in: Boissier, E. (ed.), *Flora Orientalis 1*.
Bouillé, M., & Bousquet, J. 2005. Trans-species shared polymorphisms at orthologous nuclear gene loci among distant species in the conifer *Picea* (Pinaceae): Implications for the long-term maintenance of genetic diversity in trees. *Amer. J. Bot.* 92: 63–73.
Bush, G.L. 1975. Modes of animal speciation. *Annu. Rev. Ecol. Syst.* 6: 339–364.
Cracraft, J. 1989. Speciation and its ontology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. Pp. 28-59 in: Otte, D., & Endler, J.A., (ed.), *Speciation and its consequences*. Sunderland, MA: Sinauer.
Davis, S.D., Heywood, V.H., & Hamilton, A.C. 1994. *Centres of plant diversity. A guide and strategy for their conservation. 3 Vol*. Cambridge: IUCN Publications Unit.
Engler, A. 1896. Rutaceae. Pp. 95–201 in: Engler, A., & Prantl, K., (ed.), *Nat. Pflanzenfam III, Vol. 4*. Leibzig: Engelmann.
Engler, A. 1936. Rutaceae. Pp. 187–359 in: Engler, A., & Prantl, K., (ed.), *Die natürlichen Pflanzenfamilien, 2. Aufl. 19a*. Leibzig: Engelmann.
Farris, J.S. 1969. A successive approximations approach to character weighting. *Syst. Zool.* 18: 374–385.
Farris, J.S. 1989. The retention index and the rescaled consistency index. *Cladistics* 5: 417–419.
Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the Bootstrap. *Evolution* 39: 783–791.
Fitch, W.M. 1971. Towards defining the course of evolution: minimum change for a specific tree topology. *Syst. Zool.* 20: 406–416.
Freudenstein, J.V. 1998. Paraphyly, ancestors, and classification – a response to Sosef and Brummitt. *Taxon* 47: 95–104.
Goodwillie, C., & Stiller, J.W. 2001. Evidence for polyphyly in a species of *Linanthus* (Polemoniaceae): Convergent evolution in self-fertilizing taxa. *Syst. Bot.* 26: 273–282.
Grubov, V.I. 1959. *Tentamen divisionis botanico-geographicae Asiae Centralis*. Saint Petersburg: Acad. Sci. URSS and Soc. Bot. URSS.
Harrison, R.G. 1991. Molecular changes at speciation. *Annu. Rev. Ecol. Syst.* 22: 281–308.
Hennig, W. 1966. *Phylogenetic Systematics*. Urbana: University of Illinois Press.
Hillis, D.M., & Bull, J.J. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst. Biol.* 42: 182–192.
Jalili, A., & Jamzad, Z. 1999. *Red Data Book of Iran: a preliminary survey of endemic, rare and endangered plant species in Iran*. Tehran: Research Institute of Forests and Rangelands (RIFR).
Joharchi, M. 2008. Rutaceae, vol. 60. Pp.? in: Assadi, M., Maassoumi, A.A., Babakhanlou, P., & Mozaffarian, V., (ed.), *Flora of Iran*. Tehran: Research Institute of Forests and Rangelands (RIFR).
Jussieu, A. 1825. Rutaceae: *Haplophyllum*. *Mém. Mus. Hist. Nat.* 12: 464.

Kelchner, S.A. 2000. The evolution of non-coding chloroplast DNA and its application in plant systematics. *Ann. Missouri Bot. Gard.* 87: 482–498.

Kitching, I.J., Forey, P.L., Humphries, C.J., & Williams, D.M. 1998. *Cladistics: the theory and practice of parsimony analysis*. Oxford: Oxford University Press.

Losos, J.B., & Glor, R.E. 2003. Phylogenetic comparative methods and the geography of speciation. *Trends Ecol. Evol.* 18: 220–227.

Lynch, J.D. 1989. The gauge of speciation: on the frequencies of modes of speciation. Pp. 527–553 in: Otte, D., & Endler, J.A., (ed.), *Speciation and its consequences*. Sunderland, MA: Sinauer.

Maddison, P.G., & Maddison, D.R. 2000. *MacClade4: analysis of phylogeny and character evolution*. Sunderland, MA: Sinauer.

Maddison, W.P., & Maddison, D.R. 2008. *Mesquite: a modular system for evolutionary analysis*. *Version 2.71*. http://mesquiteproject.org.

Mansion, G., Rosenbaum, G., Schoenenberger, J., Bacchetta, G., Rossello, J., & Conti, E. 2008. Invasions of the Mediterranean basin by the Araceae: integrative phylogeny-based evidence. *Syst. Biol.* 57: 269–285.

Mansion, G., Selvi, F., Guggisberg, A., & Conti, E. 2009. Origin of Mediterranean insular endemics in the Boraginales: integrative evidence from molecular dating and ancestral area reconstruction. *J. Biogeogr.* 36: 1282–1296.

Mayr, E. 1942. *Systematics and the origin of species from the viewpoint of a zoologist*. New York: Columbia Univ. Press.

Mester, I., & Vicol, E.C. 1971. Contribution to the chemotaxonomy of the *Haplophyllum* genus. *Rev. Roum. Biol.* 16: 221–233.

Moylan, E.C., Bennett, J.R., Carine, M.A., Olmstead, R.G., & Scotland, R.W. 2004. Phylogenetic relationships among *Strobilanthes* s.l. (Acanthaceae): evidence from ITS nrDNA, trnL-F cpDNA, and morphology. *Amer. J. Bot.* 91: 724–735.

Navarro, F.B., Suarez-Santiago, V.N., & Blanca, G. 2004. A new species of *Haplophyllum* A. Juss (Rutaceae) from the Iberian peninsula: evidence from morphological, karyological and molecular analyses. *Ann. Bot.* 94: 571–582.

Nazrullaev, S.S., Bessonova, I.A., & Akhmedkhodzhaeva, K.S. 2002. Estrogenic activity as a function of the chemical structure in *Haplophyllum* quinoline alkaloids. *Chem. Nat. Compd.* 37: 551–555.

Norup, M.V., Dransfield, J., Chase, M.W., Barfod, A.S., Fernando, E.S., & Baker, W.J. 2006. Homoplasious character combinations and generic delimitation: a case study from the Indo-Pacific arecoid palms (Arecaceae: Areceae). *Amer. J. Bot.* 93: 1065–1080.

Pascual-Villalobos, M.J., & Robledo, A. 1999. Anti-insect activity of plant extracts from the wild flora in southeastern Spain. *Biochem. Syst. Ecol.* 27: 1–10.

Posada, D., & Crandall, K.A. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14: 817–818.

Price, J.R. 1963. The distribution of alkaloids in the Rutaceae. Pp. 429–452 in: Swain, T., (ed.), *Chemical Plant Taxonomy*. London and New York: Academic Press.

Prieto, J.M., Giner, R.M., Recio, M.C., Schinella, G., Manez, S., & Rios, J.L. 2002. Diphyllin acetylapioside, a 5-lipoxygenase inhibitor from *Haplophyllum hispanicum*. *Planta Med.* 68: 359–360.

Pyankov, V., Black, C., Stichler, W., & Ziegler, H. 2002. Photosynthesis in *Salsola* species (Chenopodiaceae) from southern Africa relative to their C4 syndrome origin and theirAfrican-Asian arid zone migration pathways. *Plant Biology* 4: 62–69.

Quézel, P. 1978. Analysis of the flora of Mediterranean and Saharan Africa. *Ann. Missouri Bot. Gard.* 65: 479–534.

Quézel, P. 1985. Definition of the Mediterranean region and the origin of its flora. Pp. 9–24 in: Gómez-Campo, C., (ed.), *Plant conservation in the Mediterranean area*. Dordrecht: Dr. W. Junk Publishers.

Quézel, P. 1995. La flore du bassin méditerranéen: origine, mise en place, endémisme. *Ecol. Medit.* 21: 19–39.

Rambaut, A., & Drummond, A.J. 2007. *Tracer v1.4*. http://beast.bio.ed.ac.uk/Tracer.

Ribera, I., & Blasco-Zumeta, J. 1998. Biogeographical links between steppe insects in the Monegros region (Aragón, NE Spain), the eastern Mediterranean, and central Asia. *J. Biogeogr.* 25: 969–986.

Rieppel, O. 2010. Species monophyly. *J. Zool. Syst. Evol. Res.* 48: 1–8.

Ronquist, F., & Huelsenbeck, J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

Salvo, G., Bacchetta, G., Ghahremaninejad, F., & Conti, E. 2008. Phylogenetic relationships of Ruteae (Rutaceae): New evidence from the chloroplast genome and comparisons with non-molecular data. *Mol. Phylogenet. Evol.* 49: 736–748.

Sanderson, M.J., & Donoghue, M.J. 1989. Patterns of variation in levels of homoplasy. *Evolution* 43: 1781–1795.

Sang, T., Crawford, D.J., & Stuessy, T.F. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *Amer. J. Bot.* 84: 1120–1136.

Särkinen, T.S., Marcelo Peña, J.L., Yomona, A.D., Simon, M.F., Pennington, R.T., & Hughes, C.E. In press. Underestimated endemic species diversity in the Marañon seasonally dry tropical forests of Peru – an example from *Mimosa* (Leguminosae: Mimosoideae). *Taxon*.

Schierup, M.H., & Hein, J. 2000. Consequences of recombination on traditional phylogenetic analysis. *Genetics* 156: 879–891.

Scotland, R.W., Olmstead, R.G., & Bennett, J.R. 2003. Phylogeny reconstruction: the role of morphology. *Syst. Biol.* 52: 539–548.

Shaiq, M., Kashif, M., Saleem, M., & Bakhsh, R. 2001. Haplophytin-A and B: the alkaloidal constituents of *Haplophyllum acutifolium*. *Phytochemistry* 57: 1277–1280.

Shaw, J., & Small, R.L. 2005. Chloroplast DNA phylogeny and phylogeography of the North American plums (*Prunus* subgenus *Prunus* section *Prunocerasus*, Rosaceae). *Amer. J. Bot.* 92: 2011–2030.

Simmons, M. P. 2004. Independence of alignment and tree search. *Mol. Phylogenet. Evol.* 31: 874–879.

Soltani, M., & Khosravi, A.R. 2005. A new species of *Haplophyllum* from SW Iran. *Willdenowia* 35: 293–298.

Spach, E. 1949. Conspectus generis *Haplophyllum*. *Ann. Sci. Nat. Bot.* 11: 174–192.

SPSS for Windows, Rel. 11.0.1. 2001. Chicago: SPSS Inc.

Swofford, D.L. 2001. *PAUP*4.0b10: Phylogenetic Analysis Using Parsimony (*and other methods)*. Sunderland, MA: Sinauer.

Syring, J., Farrell, K., Businsk, R., Cronn, R., & Liston, A. 2007. Widespread genealogical nonmonophyly in species of *Pinus* subgenus *Strobus*. *Syst. Biol.* 56: 163–181.

Taberlet, P., Gielly, L., Pautou, G., & Bouvet, J. 1991. Universal primers for amplification of 3 non-coding regions of chloroplast DNA. *Plant Mol. Biol.* 17: 1105–1109.

Takhtajan, A. 1986. *Floristic regions of the world*. Berkeley: University of California Press.

Thompson, J.D. 2005. *Plant evolution in the Mediterranean*. Oxford: Oxford University Press.

Townsend, C.C. 1966a. Towards a revision of *Haplophyllum* A. Juss. (Rutaceae): I. *Kew Bull.* 20: 89–102.

Townsend, C.C. 1966b. *Rutaceae*. No. 36/16 in: Rechinger, K.H., (ed.), *Flora Iranica*. Graz: Akademische Druck-u. Verlagsanstalt.

Townsend, C.C. 1986. *Taxonomic revision of the genus Haplophyllum (Rutaceae)*. Hooker's Icones Plantarum, Vol. XL, parts I, II, and III. Kent, UK: Bentham-Moxon Trustees.

Venables, W.N., & Ripley, B.D. 2002. *Modern applied statistics with S. 4th ed*. New York: Springer.

Vvdensky, A. 1949. *Haplophyllum*. Pp. 200–227 in: Komarov, V.L., (ed.), *Flora of the U.R.S.S., Vol. 14*. Saint Petersburg: Academiae Scientiarum URSS.

Widmer, A., & Baltisberger, M. 1999. Extensive intraspecific chloroplast DNA (cpDNA) variation in the alpine *Draba aizoides* L. (Brassicaceae): haplotype relationships and population structure. *Molecular Ecology* 8: 1405–1415.

Wilgenbusch, J.C., Warren, D.L., & Swofford, D.L. 2004. *AWTY: A system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference*. http://ceb.csit.fsu.edu/awty.

Wilson, E.O., & Brown, W.L. 1953. The subspecies concept and its taxonomic application. *Syst. Zool.* 2: 97–111.

Yang, Z., & Rannala, B. 1997. Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo method. *Mol. Biol. Evol.* 14: 717–724.

Young, N.D., & Healy, J. 2003. Gapcoder automates the use of indel characters in phylogenetic analysis. *BMC Bioinformatics* 4: 6.

Zander, R.H. 2004. Minimal values for reliability of Bootstrap and Jackknife proportions, Decay Index, and Bayesian Posterior Probability. *PhyloInformatics* 2: 1–13.

Zohary, M. 1973. *Geobotanical foundations of the Middle East. Vol. I and II*. Stuttgart: Gustav Fischer Verlag; Amsterdam: Swets & Zeitlinger.

Figure legends

Figure 1. Left; map showing the geographic distribution of *Haplophyllum* (after Townsend, 1986) and the five floristic regions in which its species occur (after Takhtajan, 1986). Right; pie chart showing the percentage of species found in each floristic region. Note that there are no species restricted to the circumboreal floristic region only.

Figure 2. a) Fifty-percent majority rule consensus tree obtained from the Bayesian analysis on the combined molecular dataset (*mat*K, *rpl*16, *trn*K, *trn*L-*trn*F). Numbers next to branches indicate posterior probabilities (PP; > 0.50) and bootstrap percentages calculated under maximum parsimony (BP; > 50). Taxa with colored text are inferred to be non-monophyletic; taxa followed by an "E" exhibit a narrow geographical range. Colored boxes indicate the floristic region (after Takhtajan, 1986) in which each species is found (see Fig. 1 for color legend). b) Scatter plot of the first and second dimensions (x- and y-axes, respectively) of the multiple correspondence analysis of the 27 morphological characters for 45 species of *Haplophyllum* (marked with the first three letters of the species name) and its sister group *Cneoridium dumosum*.

Figure 3. Mean rescaled consistency (RC) index of each of the 27 morphological characters optimized onto 220 trees derived from the Bayesian analysis of 118 cpDNA sequences of *Haplophyllum* and outgroup taxa. Bars indicate the following categories of characters: vegetative morphology (white bars), inflorescence morphology (including features of both stamen and pistil; gray bars), and fruit morphology (black bar). Character numbers refer to those shown in Appendix 2.

Figure 4. Evolution of three selected morphological characters optimized onto the molecular phylogeny shown in Fig. 2a: "number of carpels" (Fig. 4a), "number of ovules" (Fig. 4b), and "capsule dehiscence" (Fig. 4c).

Figure 1

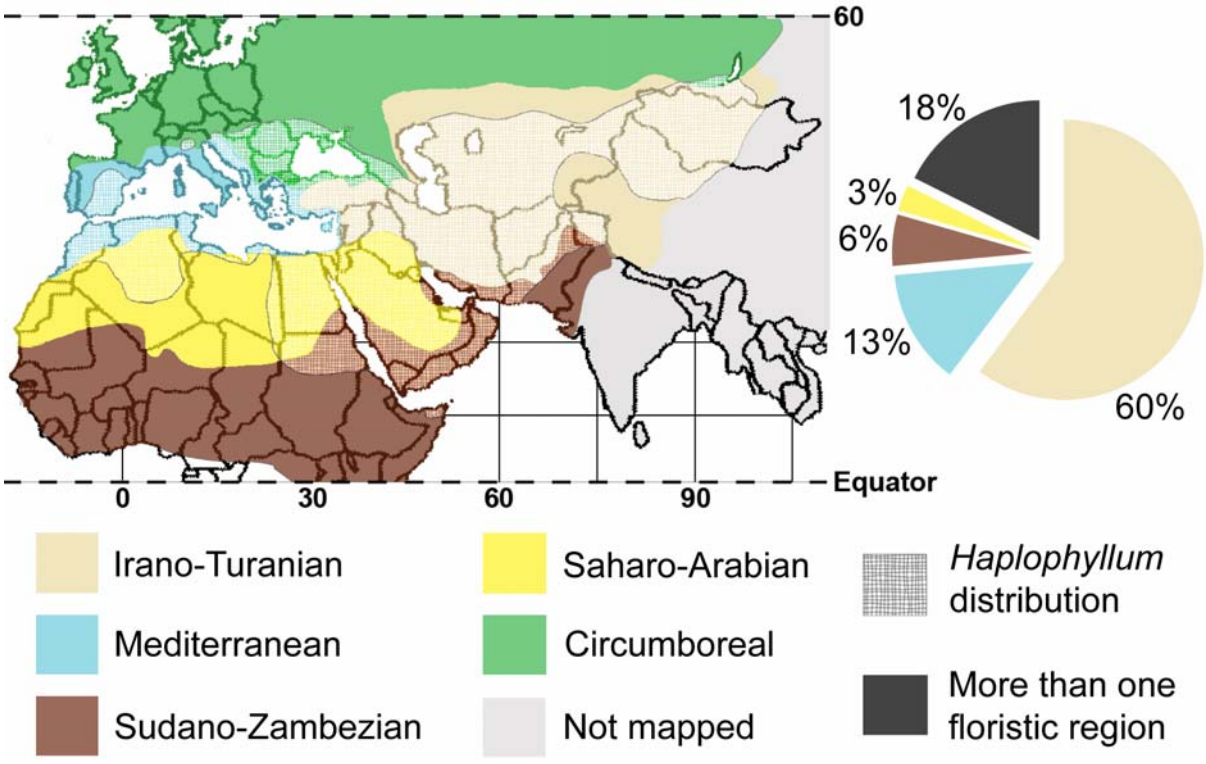

60
Equator
0
30
60
90
18%
3%
6%
13%
60%
Irano-Turanian
Saharo-Arabian
Haplophyllum distribution
Mediterranean
Circumboreal
Sudano-Zambezian
Not mapped
More than one floristic region

Figure 2

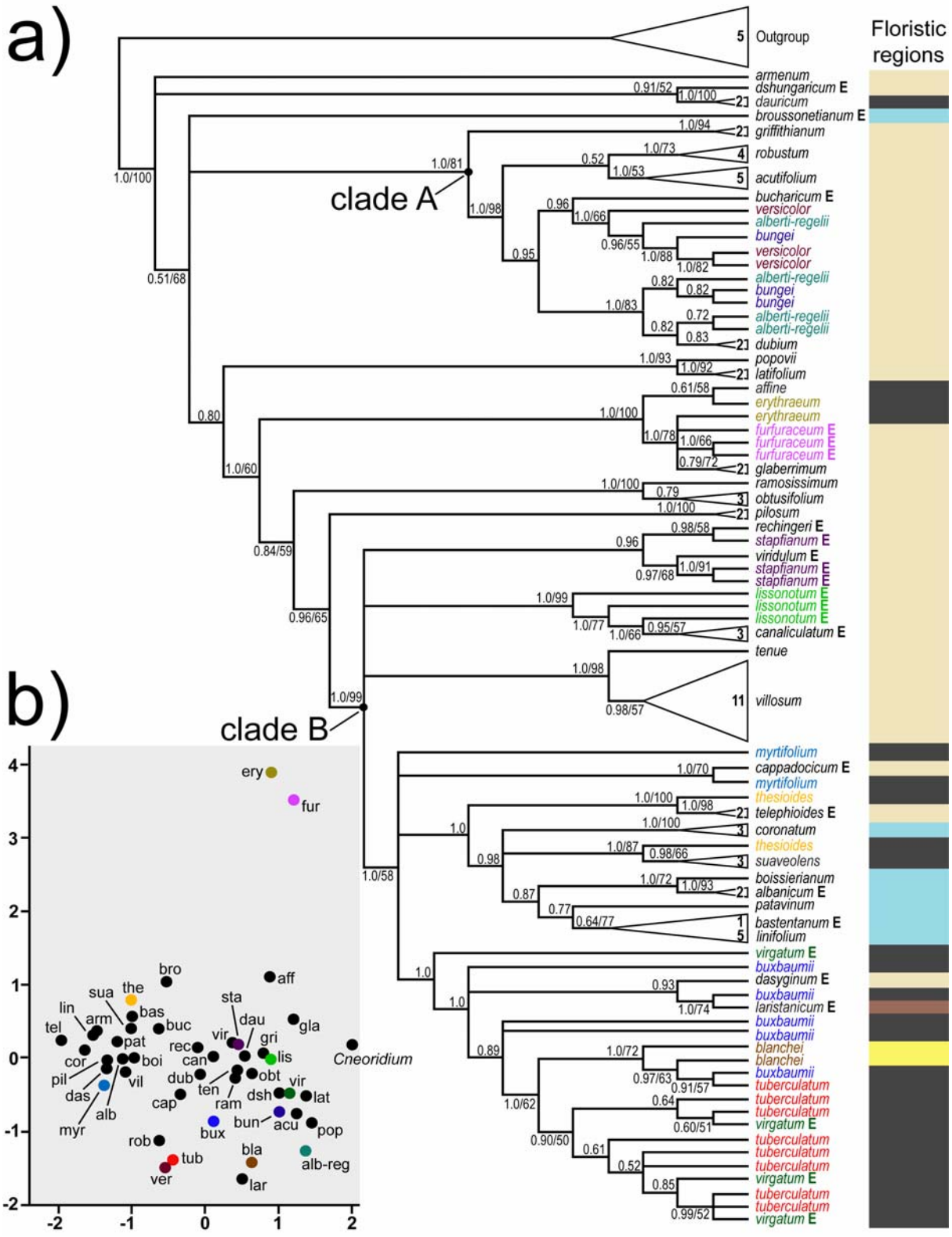

a)
b)
clade A
clade B
Floristic regions
5 Outgroup
armenum
dshungaricum E
dauricum
broussonetianum E
griffithianum
robustum
acutifolium
bucharicum E
versicolor
alberti-regelii
bungei
versicolor
versicolor
alberti-regelii
bungei
bungei
alberti-regelii
alberti-regelii
dubium
popovii
latifolium
affine
erythraeum
erythraeum
furfuraceum E
furfuraceum E
furfuraceum E
glaberrimum
ramosissimum
obtusifolium
pilosum
rechingeri E
stapfianum E
viridulum E
stapfianum E
stapfianum E
lissonotum E
lissonotum E
lissonotum E
canaliculatum E
tenue
11 villosum
myrtifolium
cappadocicum E
myrtifolium
thesioides
telephioides E
coronatum
thesioides
suaveolens
boissierianum
albanicum E
patavinum
bastentanum E
linifolium
virgatum E
buxbaumii
dasyginum E
buxbaumii
laristanicum E
buxbaumii
buxbaumii
blanchei
blanchei
buxbaumii
tuberculatum
tuberculatum
tuberculatum
virgatum E
tuberculatum
tuberculatum
tuberculatum
virgatum E
tuberculatum
tuberculatum
virgatum E
Cneoridium
ery
fur
aff
bro
sua
the
sta
lin
bas
arm
dau
gla
tel
buc
vir
gri
pat
rec
lis
cor
can
boi
obt
pil
vil
dub
vir
ten
das
dsh
lat
cap
ram
alb
myr
bux
bun
acu
pop
rob
tub
bla
alb-reg
ver
lar

Figure 3

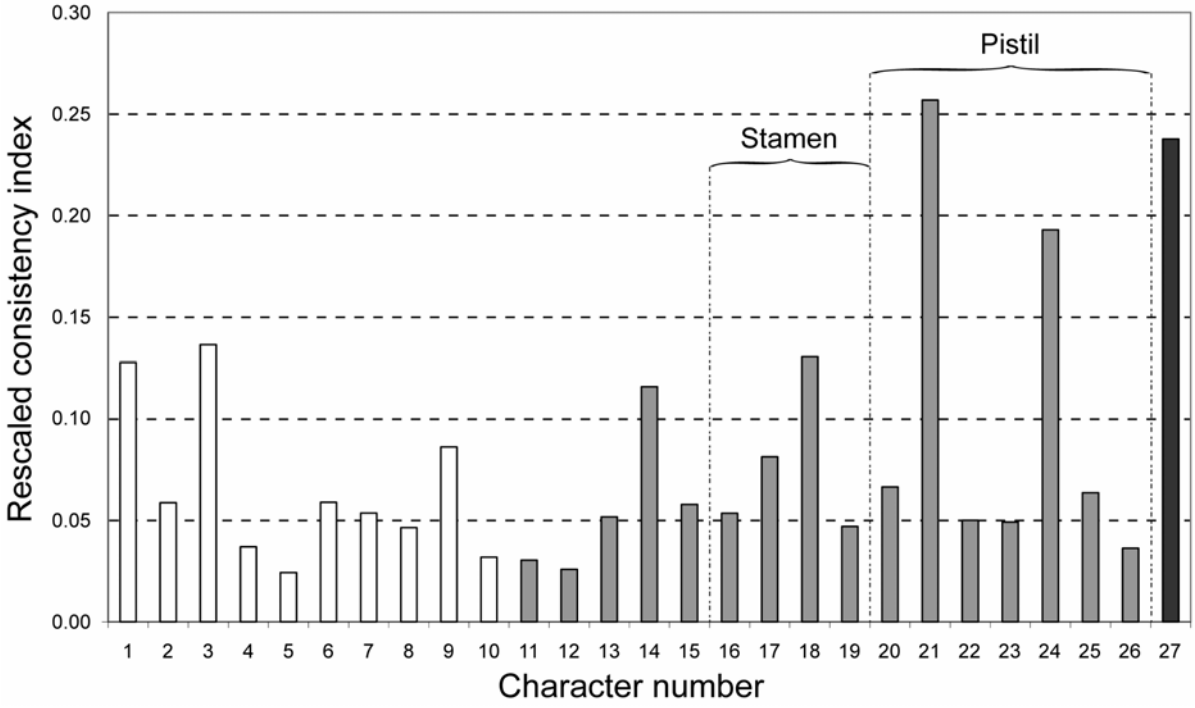

Pistil
Stamen
Rescaled consistency index
0.30
0.25
0.20
0.15
0.10
0.05
0.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
Character number

Figure 4a

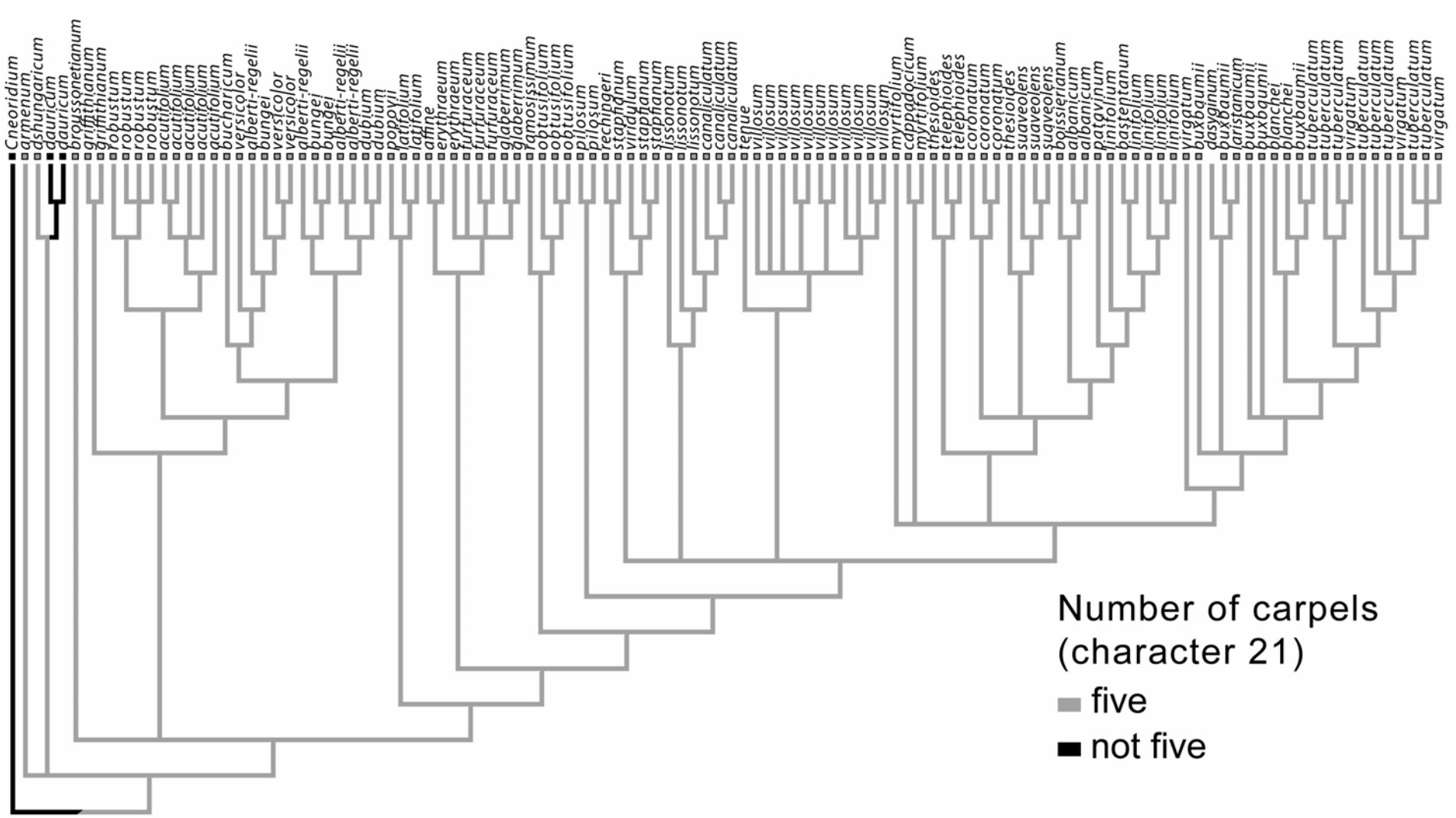

Cneoridium
armenum
ashungaricum
dauricum
dauricum
broussonetianum
griffithianum
griffithianum
robustum
robustum
robustum
robustum
acutifolium
acutifolium
acutifolium
acutifolium
acutifolium
buchariicum
versicolor
alberti-regelii
bungei
versicolor
versicolor
alberti-regelii
bungei
bungei
alberti-regelii
alberti-regelii
dubium
dubium
dubium
poposii
latifolium
latifolium
affine
erythraeum
erythraeum
turturaceum
turturaceum
turturaceum
glaberrimum
glaberrimum
ramosissimum
obtusifolium
obtusifolium
obtusifolium
pilosum
pilosum
rechingeri
staphanum
viridulum
staphanum
staphanum
lissonotum
lissonotum
lissonotum
canaliculatum
canaliculatum
canaliculatum
teque
villosum
villosum
villosum
villosum
villosum
villosum
villosum
villosum
villosum
villosum
villosum
villosum
myrtifolium
cappadocicum
myrtifolium
thesioides
telephioides
telephioides
coronatum
coronatum
coronatum
thesioides
suaveolens
suaveolens
boissieranum
albanicum
albanicum
patavinum
linifolium
bastentanum
linifolium
linifolium
linifolium
linifolium
virgatum
buxbaumii
dasyanum
buxbaumii
buxbaumii
aristanicum
buxbaumii
buxbaumii
blanchei
buxbaumii
tuberculatum
tuberculatum
tuberculatum
tuberculatum
virgatum
virgatum
tuberculatum
tuberculatum
tuberculatum
tuberculatum
virgatum
tuberculatum
tuberculatum
virgatum
Number of carpels
(character 21)
five
not five

Figure 4b

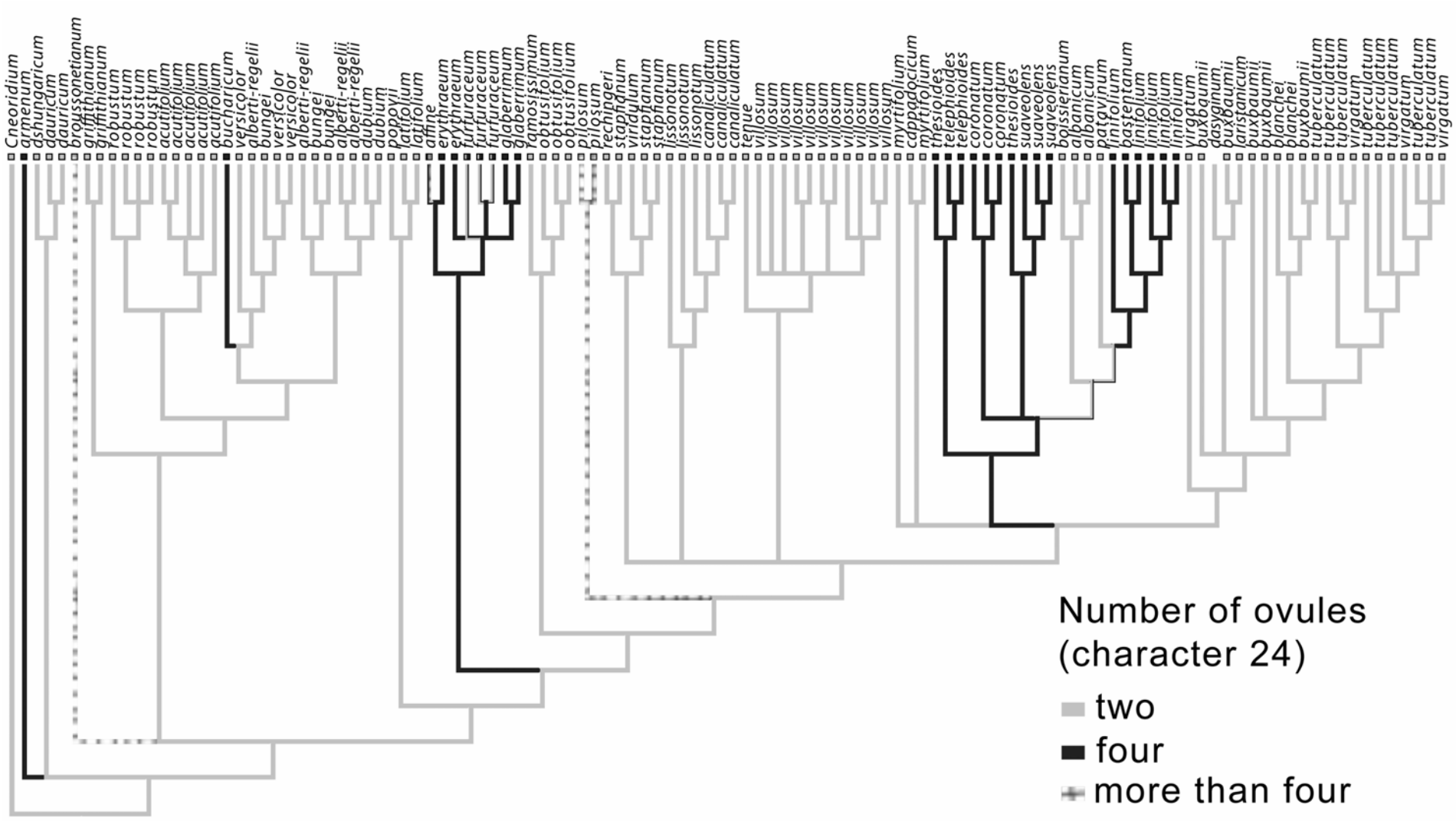

Cneoridium
armenum
ashungaricum
dauricum
dauricum
broussonetianum
griffithianum
griffithianum
robustum
robustum
robustum
robustum
acutifolium
acutifolium
acutifolium
acutifolium
acutifolium
buchariicum
versicolor
alberti-regelii
bungei
versicolor
versicolor
alberti-regelii
bungei
bungei
alberti-regelii
alberti-regelii
dubium
dubium
poppyii
latifolium
latifolium
alpine
erythraeum
erythraeum
turturaceum
turturaceum
turturaceum
glaberrimum
glaberrimum
ramosissimum
obtusifolium
obtusifolium
obtusifolium
pilosum
pilosum
rechingeri
staphanum
viridulum
staphanum
staphanum
lissonotum
lissonotum
lissonotum
canaliculatum
canaliculatum
canaliculatum
tenue
villosum
villosum
villosum
villosum
villosum
villosum
villosum
villosum
villosum
villosum
villosum
myrtifolium
myrtifolium
cappadocicum
myrtifolium
thesioides
telephioides
telephioides
coronatum
coronatum
coronatum
thesioides
suaveolens
suaveolens
suaveolens
boissierianum
albanicum
albanicum
patavinum
linifolium
bastentanum
linifolium
linifolium
linifolium
linifolium
virgatum
buxbaumii
dasyginum
buxbaumii
aristamicum
buxbaumii
buxbaumii
blanchei
blanchei
buxbaumii
tuberculatum
tuberculatum
tuberculatum
tuberculatum
virgatum
virgatum
tuberculatum
tuberculatum
tuberculatum
virgatum
tuberculatum
tuberculatum
virgatum
Number of ovules
(character 24)
two
four
more than four

Figure 4c

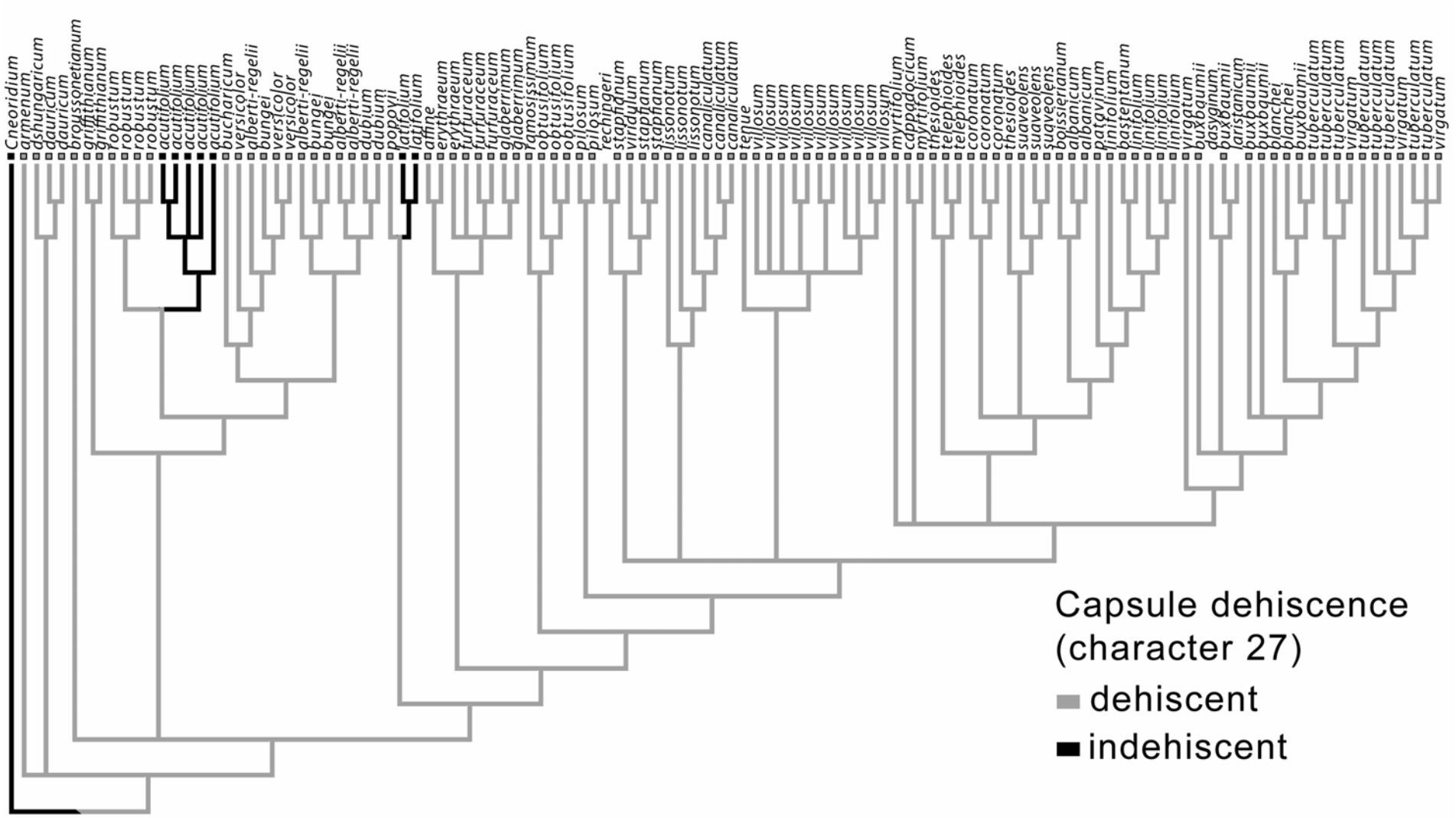

Cneoridium
armenum
ashungaricum
dauricum
dauricum
broussonetianum
griffithianum
griffithianum
robustum
robustum
robustum
robustum
acutifolium
acutifolium
acutifolium
acutifolium
acutifolium
bucharicum
versicolor
alberti-regelii
bungei
versicolor
versicolor
alberti-regelii
bungei
bungei
alberti-regelii
alberti-regelii
dubium
dubium
dubium
poposii
latifolium
latifolium
affine
erythraeum
erythraeum
turturaceum
turturaceum
turturaceum
glaberrimum
glaberrimum
ramosissimum
obtusifolium
obtusifolium
obtusifolium
pilosum
pilosum
pechipgeri
staphanum
viridulum
staphanum
staphanum
lissonotum
lissonotum
lissonotum
canaliculatum
canaliculatum
canaliculatum
teneue
villosum
villosum
villosum
villosum
villosum
villosum
villosum
villosum
villosum
villosum
villosum
villosum
myrtifolium
cappadocicum
myrtifolium
thesioides
telephioides
telephioides
coronatum
coronatum
coronatum
thesioides
suaveolens
suaveolens
suaveolens
boissierianum
albanicum
albanicum
patavinum
linifolium
bastentanum
linifolium
linifolium
linifolium
linifolium
virgatum
buxbaumii
dasyginum
buxbaumii
buxbaumii
aristanicum
buxbaumii
buxbaumii
blanchei
blanchei
buxbaumii
tuberculatum
tuberculatum
tuberculatum
tuberculatum
virgatum
virgatum
tuberculatum
tuberculatum
tuberculatum
tuberculatum
virgatum
virgatum
tuberculatum
tuberculatum
virgatum
Capsule dehiscence
(character 27)
dehiscent
indehiscent

Table 1. The two most comprehensive classifications of *Haplophyllum*.

Section	Species	Diagnostic characters
Vvedensky (1949)		
Peganoides	*H. dauricum*	Ovary: (2-)3(-4)-locular Ovules: 2 in each cell Capsule: dehiscent
Polyoon	*H. pilosum* *H. suaveolens* *H. armenum* *H. bucharicum* *H. affine*	Ovary: 5-locular Ovules: 4-12 in each cell Capsule: dehiscent
Oligoon	Remaining species (greatest bulk of the genus)	Ovary: 5-locular Ovules: 2 in each cell Capsule: dehiscent
Achaenococcum	*H. latifolium* *H. acutifolium*	Ovary: 5-locular Ovules: 2 in each cell Capsule: indehiscent
Townsend (1986)		
Peganoides	*H. gilesii* *H. dauricum*	Habit: suffrutescent perennials Flower colour: yellow or greenish-yellow Ovary: 3-locular, rarely 2- or 4-5-locular Capsule: dehiscent
Indehiscentes	*H. acutifolium* *H. latifolium*	Habit: much branched, bushy perennials Flower colour: yellow Ovary: 5-locular Capsule: indehiscent
Haplophyllum	Remaining species (greatest bulk of the genus)	Habit: Perennials, suffrutescent or herbaceous below Flower colour: white, creamy, greenish, reddish or pale to bright yellow Ovary: 5-locular Capsule: dehiscent

Table 2. Absolute number of nucleotide substitutions between selected accessions.

		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**
1	*H. blanchei*	-																				
2	*H. blanchei*	8	-																			
3	*H. buxbaumii*	21	25	-																		
4	*H. buxbaumii*	17	15	12	-																	
5	*H. buxbaumii*	16	18	17	15	-																
6	*H. buxbaumii*	2	4	14	14	10	-															
7	*H. buxbaumii*	14	12	19	9	14	11	-														
8	*H. dasyginum*	12	18	13	9	11	12	12	-													
9	*H. laristanicum*	11	14	8	13	12	4	11	10	-												
10	*H. tuberculatum*	7	11	22	18	19	6	15	15	9	-											
11	*H. tuberculatum*	8	5	15	5	11	4	4	8	9	9	-										
12	*H. tuberculatum*	10	8	20	10	15	7	7	13	9	11	0	-									
13	*H. tuberculatum*	15	13	24	14	19	12	10	17	10	17	3	3	-								
14	*H. tuberculatum*	14	18	23	19	20	14	17	16	8	15	7	8	7	-							
15	*H. tuberculatum*	12	10	21	11	16	9	9	14	9	13	2	0	5	8	-						
16	*H. tuberculatum*	12	10	21	11	16	9	9	14	9	14	1	1	3	8	2	-					
17	*H. tuberculatum*	12	10	21	11	16	9	9	14	9	14	1	1	3	8	2	0	-				
18	*H. virgatum*	20	24	23	19	20	19	18	16	14	22	14	17	23	22	20	20	20	-			
19	*H. virgatum*	14	12	23	13	18	11	11	16	10	16	2	3	1	6	4	2	2	22	-		
20	*H. virgatum*	13	11	22	12	17	10	10	15	9	14	3	0	6	9	1	3	3	21	5	-	
21	*H. virgatum*	14	12	23	13	18	11	11	16	10	16	2	3	3	8	4	2	2	22	2	5	-

Appendix 1. Sampled accessions of *Haplophyllum* and outgroup taxa, including source, voucher information, and GenBank accession numbers for the four cpDNA regions studied (sequences that could not be amplified are indicated by a "–"). Voucher specimens are deposited in the following herbaria: W = Museum of natural history of Vienna; Z = University of Zürich; LE = V. L. Komarov Botanical Institute, Saint Petersburg; TBI = Georgian Academy of Sciences, Tbilisi; GDA = University of Granada; MA = Real Jardín Botánico, Madrid; FAR = University of Tarbiat-Moaallem, Tehran; TARI = Research Institute of Forests and Rangelands, Tehran; K = Royal Botanic Gardens, Kew.

Taxon; code; source; voucher (herbarium); GenBank accession numbers: *mat*K, *rpl*16, *trn*L-*trn*F, *trn*K

Haplophyllum acutifolium* (DC.) G. Don**; acu1W; Iran, Gorgan, Golestan national park, Almeh valley; 1999-02041 (W); HM163962, HM163862, HM163761, HM163657. ***H. acutifolium; acuF1469; Iran, Qazvin Prov., Qazvin to Takestan, 14 km before Takestan; F.Ghahremaninejad 1469 (Z); EF489076, EF489150, EF489224, HM163658. ***H. acutifolium***; acuF1488; Iran, Khorassan Province, Chenaran, Freizi, 30 km S of Chenaran, 1700 m; F.Ghahremaninejad 1488 (Z); HM163963, HM163863, HM163762, HM163659. ***H. acutifolium***; acuZel0505221; Iran, Khorasan Prov., Ashkhaneh, Tange Raz; Zeltner 05.05.22 1a & b (Z); HM163964, HM163864, HM163763, HM163660. ***H. acutifolium***; acuZel0505232; Iran, Khorasan Prov., Ashkhaneh, Robat e Barah Bil; Zeltner 05.05.23 2 (Z); HM163965, HM163865, HM163764, HM163661. ***H. affine* (Aitch. & Hemsl.) Korovin**; aff1615LE; Turkmenistan; Litvinova N.P. & Nikizienko E.V. 1615 (LE); HM163966, HM163866, HM163765, HM163662. ***H. albanicum* (Bald.) Bornm.**; alb85697LE; Macedonia; E. Mayer 85697 (LE); HM163967, HM163867, HM163766, HM163663. ***H. albanicum***; albSelvi5; Albania, Drisht, Scutari region; Selvi, Coppi, Cecchi 5 (Z); HM163968, HM163868, HM163767, HM163664. ***H. alberti-regelii* Korovin**; albert220LE; Tajikistan; V.P. Bochantzev 220 (LE); HM163969, HM163869, HM163768, HM163665. ***H. alberti-regelii***; albert2W; Afghanistan, Baghlan, Surkh Kotal, ca. 15km NW of Pule-Khumri; 1976-00031 (W); HM163970, –, HM163769, HM163666. ***H. alberti-regelii***; albert391LE; Uzbekistan; V.P. Bochantzev 391 (LE); HM163971, HM163870, HM163770, –. ***H. alberti-regelii***; albert3W; Afghanistan, Badakhshan, 15 miles NE of Kesem, road to Faizabad; 1973-13349 (W); HM163972, HM163871, HM163771, HM163667. ***H. armenum* Spach**; armeT13; Georgia; s.n. (TBI); HM163973, HM163872, HM163772, HM163668. ***H. bastentanum* F.B. Navarro, Suár.-Sant. & Blanca**; bastGDA47502; Spain; 47502 (GDA); EF489097, EF489171, EF489245, HM163669. ***H. blanchei* Boiss.**; blan6W; Iraq, desertum occidentale, inter Ramadi et Rutba 260 km; 16372 (W); HM163974; HM163873; HM163773; HM163670. ***H. blanchei***; blan7W; Jordanien, Amman, Nordostjordanische Basaltwuste, Hammada, ca. 50 km W of Azraq; 2004-20318(W); HM163975, HM163874, HM163774, HM163671. ***H. boissierianum* Vis. & Pančič**; boissSelvi2; Albania, Krume, Mt. Pastrik, Region of Kukes; Selvi, Coppi, Cecchi 2 (Z); HM163976, HM163875, HM163775, HM163672. ***H. broussonetianum* Coss.**; broussW; Marokko, Todra-schlucht; 1999-06842(W); HM163977, HM163876, HM163776, HM163673. ***H. bucharicum* Litv.**; buch211Z; Uzbekistan, betw. Shurab and Darhand; Manafzadeh & Salvo 211 (Z); HM163978, HM163877, HM163777, HM163674. ***H. bungei* Trautv.**; bun207Z; Uzbekistan, Shafrikan-Shuruk village, 51 km after Shafrikan (Botanical desert station), SW Kizil Kum; Manafzadeh & Salvo 207 (Z); HM163979, HM163878, HM163778, HM163675. ***H. bungei***; bun2119LE; Uzbekistan; R.V. Kamelin 2119 (LE); HM163980, HM163879, HM163779, HM163676. ***H. bungei***; bun431LE; Kazakhstan, western part; I.N. Saffronova et al. 431 (LE); –, HM163880, HM163780, –. ***H. buxbaumii* (Poir.) G. Don**; bux12W; Iraq, Hamam Ali; 1974-06575 (W); HM163981, HM163881, HM163781, HM163677. ***H. buxbaumii***; bux13W; Iraq, Rasheed; 1970-1918 (W); HM163982, HM163882, HM163782, HM163678. ***H.***

buxbaumii; bux14W; Turkey, 3 km S of Caykavak pass C5, Nigde; 1991-9560 (W); HM163983, HM163883, HM163783, HM163679. ***H. buxbaumii***; buxMA557457; Tunisia; 557457 (MA); EF489095, EF489169, EF489243, HM163680. ***H. buxbaumii***; buxTurkey; Turkey, 1km before Nizip; Gabriele 19 May 2006 b (Z); HM163984, HM163884, HM163784, HM163681. ***H. canaliculatum* Boiss.**; canF1454; Iran, Fars Prov., Shiraz to Kharameh, km 13; F.Ghahremaninejad 1454 (Z); EF489077, EF489151, EF489225, HM163682. ***H. canaliculatum***; canSM50; Iran, Fars Prov., between Bidshahr and Kavian, near Banarooye; Manafzadeh 5 (Z); HM163985, HM163885, HM163785, HM163683. ***H. canaliculatum***; canZel0505092; Iran, Paste chenar region, 10 km SE of Sarvestan; Zeltner 05.05.09 2 (Z); HM163986, HM163886, HM163786, HM163684. ***H. cappadocicum* Spach**; cap16W; Turkey, Eski Malata; 1965-19219 (W); HM163987, HM163887, HM163787, HM163685. ***H. cappadocicum***; cor17W; Turkey; 1970-16869 (W); HM163988, HM163888, HM163788, HM163686. ***H. coronatum* Griseb.**; cor19W; Greece, Thessalien, ca. 28 km WNW of Kalambaka; 2000-14981 (W); HM163989, HM163889, HM163789, HM163687. ***H. coronatum***; corMA353234; Macedonia; 353234 (MA); EF489081, EF489155, EF489229, HM163688. ***H. dasyginum* C. Towns.**; dasyW; Iran, Hamadan Prov. Aq Bulaq; 1961-1137 (W); HM163990, HM163890, HM163790, HM163689. ***H. dauricum* (L.) G. Don**; daur354LE; Mongolia; V.I. Grubov et al. 354 (LE); HM163991, HM163891, HM163791, HM163690. ***H. dauricum*** ; daurOyumaa1; Mongolia; Oyumaa 1 (Z); EF489099, EF489173, EF489247, HM163691. ***H. dshungaricum* Rubtzov**; dshun616LE; Kazakhstan, mountains in the eastern part; M. Piminov et al. 616 (LE); HM163992, HM163892, HM163792, HM163692. ***H. dubium* Korovin**; dub22W; Afghanistan, Faryab province, 7 miles E of Maimana, road to Belcheragh; 1973-13418 (W); HM163993, HM163893, HM163793, HM163693. ***H. dubium***; dub49LE; Turkmenistan; V.P. Bochantzev 49 (LE); HM163994, HM163894, HM163794, HM163694. ***H. erythraeum* Boiss.**; ery24W; Afghanistan, Farah province, 21.5 miles E of Dilaram, road to Kandahar; 1973-13384(W); HM163995, HM163895, HM163795, HM163695. ***H. erythraeum***; eryZel0505101; Iran, Fars Prov., Reserve de Onagres; Zeltner 05.05.10 1 (Z); HM163996, HM163896, HM163796, HM163696. ***H. furfuraceum* Bunge ex Boiss.**; fur104Z; Iran, Shahrood-Ramiyan road, near to military campus, 85 km to Azadshahr; Manafzadeh & Salvo 104 (Z); HM163997, HM163897, HM163797, HM163697. ***H. furfuraceum***; fur26W; Iran, C Damghan-Semnan, in deserto gypsaceo, 2-7 km supra Sorkheh, prope Semnan; 1983-07483 (W); HM163998, HM163898, HM163798, HM163698. ***H. furfuraceum***; furFAR14508; Iran, Khorasan Prov., between Mashad & Sarakhs, Chaahak hills; 14508 (FAR); EF489093, EF489167, EF489241, HM163699. ***H. glaberrimum* Bunge ex Boiss.**; gla28W; Iran, C. Kavir (Kavir protected region), Mobarakiyeh 40 km a Varamin, eridiem versus; 1975-13435 (W); HM163999, HM163899, HM163799, HM163700. ***H. glaberrimum***; glaFAR34555; Iran, Khorasan Prov., SW of Sabz e vaar, Parvand; 34555 (FAR); EF489082, EF489156, EF489230, HM163701. ***H. griffithianum* Boiss.**; gri182LE; Tajikistan; V.P. Bochantzev 182 (LE); –, HM163900, HM163800, –. ***H. griffithianum***; gri30W; Afghanistan, Kalifghan, Kataghan Prov.; 1980-16250 (W); HM164000, HM163901, HM163801, HM163702. ***H. laristanicum* C. Towns.**; lariW; Iran, Lar prov.; 1958-2917 (W); –, HM163902, HM163802, –. ***H. latifolium* Kar. & Kir.**; lat1971LE; Kazakhstan, southern part; R.V. Kamelin 1971 (LE); HM164001, HM163903, HM163803, HM163703. ***H. latifolium***; latMA642325; Uzbekistan; 642325 (MA); EF489094, EF489168, EF489242, HM163704. ***H. linifolium* (L.) G. Don**; lin11758LE; Marocco; J. Lewalle 11758 (LE); HM164002, HM163904, HM163804, –. ***H. linifolium***; lin31W; Spain, Prov. Huesca, in collibus siccis 10 km a candasnos meridiem versus, subste. Calcif; 1994-09998 (W); HM164003, HM163905, HM163805, HM163705. ***H. linifolium***; lin32W; Spain, Madrid, entre Aranjuez y Valdelagua. Base del cwrro cavina; 1995-01052 (W); HM164004, HM163906, HM163806, HM163706. ***H. linifolium***; linGDA47314; Spain; 47314 (GDA); HM164005, HM163907, HM163807, HM163707. ***H. linifolium***; linMA684635; Spain, Almeria; 684635 (MA); EF489079, EF489153, EF489227,

HM163708. ***H. lissonotum* C. Towns.**; lisSM30; Iran, Hormozgan Prov., Bastak-Lar road (5 km after Bastak); Manafzadeh 3 (Z); HM164006, HM163908, HM163808, HM163709. ***H. lissonotum***; lisSM40; Iran, Hormozgan Prov., Bastak-Lar road (8km after Bastak); Manafzadeh 4 (Z); HM164007, HM163909, HM163809, HM163710. ***H. lissonotum***; lisZel0505112; Iran, Hormozgan Prov., Mount. Genu; Zeltner 05.05.11 2 & 3 (Z); HM164008, HM163910, HM163810, HM163711. ***H. myrtifolium* Boiss.**; myr36W; Turkey, Malayta, Yeshilyurt; miozaner mergel; 1965-19217 (W); HM164009, HM163911, HM163811, HM163712. ***H. myrtifolium***; myr37W; Turkey, distr. Asaray, in valle ihlara; 2004-07473 (W); HM164010, HM163912, HM163812, HM163713. ***H. obtusifolium* (Ledeb.) Ledeb.**; obt106Z; Iran, Ghoochan-Dargaz, 22 km to Dargaz, Bibi gherghez; Manafzadeh & Salvo 106 (Z); HM164011, HM163913, HM163813, HM163714. ***H. obtusifolium***; obt38W; Iran, NE, ca. 7 km NW of Soolegerd; 1999-02467 (W); –, HM163914, HM163814, HM163715. ***H. obtusifolium***; obtFAR16876; Iran, Khorasan Prov., South of Daragaz, Gherkhghez hills; 16876 (FAR); EF489098, EF489172, EF489246, –. ***H. patavinum* (L.) G. Don**; patMA353204; Bosnia and Herzegovina; 353204 (MA); EF489085, EF489159, EF489233, –. ***H. pilosum* Stschegleev ex Turcz.**; pil204Z; Uzbekistan, border between Kattakurgan and past Kargan; Manafzadeh & Salvo 204 (Z); HM164012, HM163915, HM163815, HM163716. ***H. pilosum***; pilFAR18330; Iran, Khorasan Prov., between Torbat e Heydarieh & Gonabad, Lut e Omrani, Ziarat; 18330 (FAR); HM164013, HM163916, HM163816, HM163717. ***H. popovii* Korovin**; popo44W; Afghanistan, E. khosht, in montibus S yakubi, substr. Kalkschiefer; 1968-202 (W); –, HM163917, HM163817, –. ***H. ramosissimum* (Paulsen) Vved.**; ram1968LE; Uzbekistan, Karakalpakstan; Sherbaev 1968 (LE); –, HM163918, HM163818, –. ***H. rechingeri* C. Towns.**; rechSM90; Iran, Bakhtiari Prov., Boroojen to Sefiddasht, 2km to Zarrinshahr junction; Manafzadeh 9 (Z); HM164014, HM163919, HM163819, –. ***H. robustum* Bunge**; rob100Z; Iran, Tehran-Semnan, south of Tehran, 6km to Eyvanaki; Manafzadeh & Salvo 100 (Z); HM164015, HM163920, HM163820, HM163718. ***H. robustum***; rob108Z; Iran, Kashan, Aran-Bidgol, Maranjab area, 5 km after Aran; Manafzadeh & Salvo 108 (Z); HM164016, HM163921, HM163821, HM163719. ***H. robustum*** ; rob206Z; Uzbekistan, Navai, 30 km after Navai- Shafreghan, M-37 (Tashkent- Buchara); Manafzadeh & Salvo 206 (Z); HM164017, HM163922, HM163822, HM163720. ***H. robustum*** ; robFAR17725; Iran, Khorasan Prov., between Torbat e Heydarieh & Gonabad, Lut e Omrani; 17725 (FAR); EF489089, EF489163, EF489237, HM163721. ***H. stapfianum* Hand.-Mazz.**; stapfianumSM20; Iran, Kerman Prov., Sirjan-Bandar road (270km to Bandar); Manafzadeh 2 (Z); HM164018, HM163923, HM163823, HM163722. ***H. stapfianum***; staF1426; Iran, Fars Prov., Shiraz to Kharameh, km 21; F.Ghahremaninejad 1426 (Z); EF489078, EF489152, EF489226, HM163723. ***H. stapfianum***; staF1448; Iran, Fars Prov., beginning of the road Shiraz to Zarghan; F.Ghahremaninejad 1448 (Z); HM164019, HM163924, HM163824, HM163724. ***H. suaveolens* (DC.) G. Don**; sua49W; Bulgaria, Rousse region, Danube plain; 2005-05940 (W); –, HM163925, HM163825, HM163725. ***H. suaveolens***; sua63W; Bulgaria, Central rhodope mts., above the town of Assenovgrad; 2007-05771 (W); HM164020, HM163926, HM163826, –. ***H. suaveolens***; suaMA692105; Macedonia; 692105 (MA); EF489086, EF489160, EF489234, HM163726. ***H. telephioides* Boiss.**; tel50W; Turkey, E Ozkonak, B5 Nevshehir; 1990-06922 (W); HM164021, HM163927, HM163827, HM163727. ***H. telephioides***; telTurkey; Turkey, Yaylaci; Gabriele 21 May 2006 b (Z); HM164022, HM163928, HM163828, HM163728. ***H. tenue* Boiss.**; ten51W; Armenia, vayots dzor province, eghegnadzor district, E part; 2006-03911 (W); HM164023, HM163929, HM163829, HM163729. ***H. thesioides* (Fisch. ex DC.) G. Don**; thes52W; Turkey, 22 km W Tarakli; 1990- 07118 (W); HM164024, HM163930, HM163830, HM163730. ***H. thesioides***; thesTurkey; Turkey, Bozgüney; Gabriele 21 May 2006 x (Z); HM164025, HM163931, HM163831, HM163731. ***H. tuberculatum* (Forssk.) Adr. Juss.**; tub60W; Egypt, NW/C part of the oasis 6-10 km NNW of Mt…; 2006-12053 (W); HM164026, HM163932, HM163832,

HM163732. ***H. tuberculatum***; tub61W;Tunisia, SE, ca. 35 km E Medenine; 2007-07253 (W); –, HM163933, HM163833, HM163733. ***H. tuberculatum***; tubTARI72376; Iran, Semnan Prov., Sorkheh, Lasjerd; 72376 (TARI); HM164027, HM163934, HM163834, –. ***H. tuberculatum*** ; tubZel0505081a; Iran, Fars Prov., Firuzabad, Palace of Ardeshir's I; Zeltner 05.05.08 1a (Z); HM164028, HM163935, HM163835, HM163734. ***H. tuberculatum***; tubZel0505081b; Iran, Fars Prov., Firuzabad, Palace of Ardeshir's I; Zeltner 05.05.08 1 (Z); HM164029, HM163936, HM163836, HM163735. ***H. tuberculatum***; tubZel0505152; Iran, Kerman Prov., Shahdad; Zeltner 05.05.15 2a & b (Z); HM164030, HM163937, HM163837, HM163736. ***H. tuberculatum***; tubZel1; Iran, Lali, 5 km avant; Zeltner & Mansion 3.5.2007 sn1 (Z); HM164031, HM163938, HM163838, HM163737. ***H. tuberculatum***; tubZel2; Iran, Choghazanbil; Zeltner & Mansion 3.5.2007 sn2 (Z); HM164032, HM163939, HM163839, HM163738. ***H. versicolor*** **Fisch. & Mey.**; ver203Z; Uzbekistan, 6 km before Samarkand; Manafzadeh & Salvo 203 (Z); HM164033, HM163940, HM163840, HM163739. ***H. versicolor***; ver209Z; Uzbekistan, border between Guzar and Dekanabad regions, 15km from Pachkamar village, coming from Guzar; Manafzadeh & Salvo 209 (Z); HM164034, HM163941, HM163841, HM163740. ***H. versicolor***; verFAR23840; Iran, Khorasan Prov., Sarakhs, between Doulatabad and Polekhatoon; 23840 (FAR); EF489090, EF489164, EF489238, HM163741. ***H. villosum*** **(M. Bieb.) G. Don**; vil59W; Georgia, E, ca. 40 km SE of Tziteli-Tzkaro, S limit of Shirak plateau, Vashlovan valley, Pantishar gorge; 2008-02417 (W); –, HM163942, HM163842, HM163742. ***H. villosum***; vil9LE; Caucasus region; Menitzki & Popova 9 (LE); –, HM163943, HM163843, –. ***H. villosum***; vilMA417870; Azerbaijan; 417870 (MA); EF489096, EF489170, EF489244 , –. ***H. villosum***; vilSM110; Iran, East Azarbayjan Prov., Khaje to Ahar road (6km after Khaje); Manafzadeh 11 (Z); HM164035, HM163944, HM163844, HM163743. ***H. villosum***; vilSM120; Iran, East Azarbayjan Prov., Kaleybar (Orliban Dam); Manafzadeh 12 (Z); HM164036, HM163945, HM163845, HM163744. ***H. villosum***; vilSM130; Iran, East Azarbayjan Prov., Kaleybar (Orliban Dam); Manafzadeh 13 (Z); HM164037, HM163946, HM163846, HM163745. ***H. villosum***; vilSM140; Iran, East Azarbayjan Prov., Bandar e Sharafkhaneh-Tasooj road, Heris Village; Manafzadeh 14 (Z); HM164038, HM163947, HM163847, HM163746. ***H. villosum***; vilSM150; Iran, East Azarbayjan Prov., Sarab to Heris Village, after Asbforooshan junction; Manafzadeh 15(Z); HM164039, HM163948, HM163848, HM163747. ***H. villosum***; vilSM160; Iran, Ardebil Prov., Ardebil-Nir road (12 km to Nir); Manafzadeh 16 (Z); HM164040, HM163949, HM163849, HM163748. ***H. villosum***; vilSM170; Iran, Ardebil Prov., Ardebil-Nir road (12 km to Nir); Manafzadeh 17(Z); HM164041, HM163950, HM163850, HM163749. ***H. villosum***; vilT; Georgia, Transcaucasia; s.n. (TBI); HM164042, HM163951, HM163851, HM163750. ***H. virgatum*** **Spach**; virgF1428; Iran, Fars Prov., Shiraz to Kharameh, km 21; F.Ghahremaninejad 1428 (Z); HM164043, HM163952, HM163852, HM163751. ***H. virgatum***; virgF1437; Iran, Fars Prov., Shiraz to Kazerrun, Parishan Lake; F.Ghahremaninejad 1437 (Z); HM164044, HM163953, HM163853, HM163752. ***H. virgatum***; virgSM60; Iran, Fars Prov., Jahrom-Shiraz road (85km to Shiraz); Manafzadeh 6 (Z); HM164045, HM163954, HM163854, HM163753. ***H. virgatum***; virgSM70; Iran, Fars Prov., Dahak Village, 100 km to solar Powerhouse; Manafzadeh 7 (Z); HM164046, HM163955, HM163855, HM163754. ***H. viridulum*** **Soják**; virSM80; Iran, Fars Prov., Shiraz-Fasa road (Miyanjangal), opposite of emamzadeh Esmail; Manafzadeh 8 (Z); HM164047, HM163956, HM163856, HM163755. ***Aegle marmelos*** **Corrêa**; Aeg; Eastern Asia; Chase 1340 (K); HM163957, HM163857, HM163756, HM163653. ***Citrus reticulata*** **Blanco**; Citr; Switzerland, Zürich Botanic Gardens, living collection, cult. 19790418; Sandro Wagen 48 (Z); HM163958, HM163858, HM163757, –. ***Cneoridium dumosum*** **Hook. f.**; Cneo; USA, Oak Crest Park, California; Alexander Kocyan 154 (Z); HM163959, HM163859, HM163758, HM163654. ***Poncirus trifoliata*** **(L.) Raf.**; Ponc; Switzerland, Zürich Botanic Gardens, living collection, cult. 19760414; Sandro Wagen 7 (Z); HM163960, HM163860, HM163759,

HM163655. ***Glycosmis citrifolia* Lindl.**; Glyc; Taiwan, Taipei; Yih-Han Chang 3310 (Z); HM163961, HM163861, HM163760, HM163656.

Appendix 2. Morphological characters and states selected for this study.

Vegetative morphology

1. Number of stems: (1) one, (2) more than one.
2. Inflorescence form in each stem: (1) lax/broad, (2) dense/compact.
3. Sterile axillary shoots: (0) absent, (1) present.
4. Glands on stem: (0) invisible under a microscope (magnification 40x), (1) visible.
5. Stem branching: (1) branched under the inflorescence, (2) unbranched.
6. Stem indumentum: (0) glabrous, (1) hairy scattered, (2) hairy dense.
7. Indumentum of leaf margin: (0) glabrous, (1) hairy.
8. Indumentum of leaf: (0) glabrous, (1) hairy.
9. Tuberculate glands on leaf: (0) absent, (1) present.
10. Petiole: (0) absent, (1) present.

Inflorescence morphology

11. Indumentum of inflorescence: (0) glabrous, (1) hairy.
12. Form of bract: (1) linear, (2) broad.
13. Indumentum of sepal: (0) glabrous, (1) farinose, (2) hairy.
14. Dark dorsal vitta/tinge on petal: (0) absent, (1) present.
15. Indumentum of petal: (0) glabrous, (1) farinose, (2) hairy.

Stamen

16. Form of filament: (1) abruptly expanding from base to apex, (2) gradually expanding.
17. Attachment between filaments: (1) monadelphous, (2) free.
18. Indumentum of filament: (0) glabrous, (1) hairy in central portion, (2) hairy in lower half.
19. Form of anther: (1) oval, (2) oblong.

Pistil

20. Apical appendage on ovary: (0) absent, (1) present.
21. Number of carpels: (1) five, (2) not five.
22. Indumentum of ovary: (0) glabrous, (1) farinose, (2) hairy.
23. Glands of ovary: (0) non tuberculate, (1) tuberculate.
24. Number of ovules: (2) two, (4) four, (5) more than four.
25. Form of style: (1) slender, (2) stout.
26. Indumentum of style: (0) glabrous, (1) hairy.

Fruit morphology

27. Capsule dehiscence: (1) dehiscent, (2) indehiscent.

Appendix 3. Morphological matrix for the 45 species of *Haplophyllum* and its sister species, *Cneoridium dumosum*, included in this study. Character numbers refer to those presented in Appendix 2. Missing data are indicated by a "?"; polymorphic character states are indicated with a "&".

Taxon/Character	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Haplophyllum acutifolium	2	1	0	1	1	0	1	0	0	1	1	2	0	0	0	1	2	1	1
H. affine	2	2	1	1	2	0	0	0	0	0	0	2	0	0	0	2	2	0	2
H. albanicum	2	2	1	1	2	2	?	?	0	1	1	1	2	0	0	2	2	2	1
H. alberti-regelii	1	1	?	1	1	0	0	0	0	0	0	1	2	0	0	1	1	1	2
H. armenum	2	2	1	0	2	2	0	0	0	1	1	1	2	1	2	2	2	1	1&2
H. bastentanum	2	2	1	?	?	2	0	0	0	0	1	1	2	0	2	2	2	2	2
H. blanchei	2	1	1	1	1	2	1	1	0&1	0	1	1	0	0	0	1	1	1	1&2
H. boissierianum	2	1	1	1	2	2	0	0	0	1	1	1	2	1	0	2	2	2	1
H. broussonetianum	2	2	1	1	?	1	0	0	0	1	0&1	1	2	0	0	2	2	0	1
H. bucharicum	2	1	1	0	2	1	1	0	0	?	1	1	2	0	0	2	2	2	2
H. bungei	2	1	1	1	1	0	0	0	0	1	0	2	0	0	0	1	1	1	2
H. buxbaumii	2	1	1	1	2	?	1	1	0	1	1	2	0	0	0	1	2	1	1
H. canaliculatum	2	2	1	1	2	0	0	0	0	1	1	2	0	0	0	2	2	1	2
H. cappadocicum	2	1	1	0	1	2	1	0	0	1	1	1	2	0	0	2	2	1	2
H. coronatum	2	2	1	0	2	?	1	1	0	1	1	1	2	0	2	2	2	2	2
H. dasygynum	2	2	?	1	2	1	1	1	0	?	1	1	2	0	0	?	?	?	?
H. dauricum	2	2	1	1	2	?	0	0	0	0	1	2	2	0	0	1	2	2	2
H. dshungaricum	2	1	1	1	2	0	0	0	0	?	0	1	0	0	0	1	2	1	2
H. dubium	2	1	1	0	2	0	0	0	0	1	1	1	2	0	0	1	1&2	2	2
H. erythraeum	2	2	1	1	1	0	0	0	0	0	1	2	1	0	1	2	2	0	2
H. furfuraceum	2	1	1	1	2	0	0	0	0	0	1	1	1	0	1	1	2	0	2
H. glaberrimum	2	1	1	1	1	0	0	0	0	0	0	2	0	0	0	2	2	0	2
H. griffithianum	2	1&2	1	1	2	0	0	0	0	0	0	1	0	0	0	2	2	1	1&2
H. laristanicum	?	1	?	0	1	?	1	1	0	0	1	?	2	0	0	1	1	1	2
H. latifolium	2	1	0	1	2	0	0	0	0	1	0	2	0	0	0	1	2	1	2
H. linifolium	2	2	1	1	2	1	1	1	0	0	1	1	2	1	0	2	2	2	2
H. lissonotum	2	1	1	1	2	0	0	0	0	0	0	1	0	0	0	2	2	1	2
H. myrtifolium	2	2	1	0	2	1	1	1	0	0	1	2	2	0	2	2	2	1	1
H. obtusifolium	2	1	0	1	2	2	0	0	1	0	0	2	0	0	0	2	2	1	2
H. patavinum	2	2	1	0	2	2	0	0	0	1	1	1	2	1	2	2	2	2	2
H. pilosum	2	2	0	1	2	2	1	1	0	1	1	1	2	1	0	2	2	1	1&2
H. popovii	2	1	0	1	1	0	0	0	0	1	0	1	0	0	0	1	1&2	1	2
H. ramosissimum	2	1	?	0	1	?	0	0	1	0	1	1	0	0	0	2	2	1	1&2
H. rechingeri	2	2	1	1	2	0	0	0	0	1	1	1	2	0	2	2	2	1	2
H. robustum	1	1	0	1	2	?	1	1	1	1	1	1	2	0	2	1&2	2	2	2
H. stapfianum	2	2	1	1	2	0	0	0	0	1	0	1	0	0	0	2	2	2	2
H. suaveolens	2	2	1	1	2	2	1	0	0	0	1	2	0	1	2	2	2	1	2
H. telephioides	2	2	1	0	2	?	1	1	0	1	1	2	2	1	2	2	2	2	1
H. tenue	2	2	0	1	2	0	0	0	0	1	1	2	2	0	0	1	2	1	2
H. thesioides	2	2	1	1	2	1	1	0	0	0	1	2	2	1	0	2	2	2	2
H. tuberculatum	2	1	1	1	2	1	1	1	1	0	1	1&2	2	0	0	2	1	1	2
H. versicolor	2	1	?	1	1	2	1	1	0	1	1	1	2	0	2	1	1	1	2
H. villosum	2	2	1	1	2	?	1	1	0	1	1	1	2	0	0	2	2	2	2
H. virgatum	2	1	1	1	1	0	0	0	0	0	1	1	0	0	0	1	2	1	2
H. viridulum	2	1	1	1	2	0	0	0	0	0	0	2	2	0	2	2	2	2	2
Cneoridium dumosum	2	1	?	1	?	0	0	0	0	0	0	?	0	0	0	2	2	0	1&2

Continued.

Taxon/Character	20	21	22	23	24	25	26	27
Haplophyllum acutifolium	0	1	0	0	2	2	0	2
H. affine	1	1	0	0	4&5	2	0	1
H. albanicum	1	1	2	1	2	1	0	1
H. alberti-regelii	0	1	0	0	2	1	0	1
H. armenum	1	1	2	1	4	1	0	1
H. bastentanum	?	1	2	?	4	1	0	1
H. blanchei	0	1	0	0	2	1	0	1
H. boissierianum	1	1	2	1	2	1	0	1
H. broussonetianum	1	1	0	1	5	2	0	1
H. bucharicum	1	1	0	0	4	2	0	1
H. bungei	1	1	0	1	2	1	0	1
H. buxbaumii	0	1	2	0	2	1	0	1
H. canaliculatum	0	1	2	0	2	2	0	1
H. cappadocicum	0	1	2	1	2	2	0	1
H. coronatum	1	1	2	1	4	1	0	1
H. dasygynum	?	?	?	?	?	?	?	?
H. dauricum	0	2	0	0	2	1	0	1
H. dshungaricum	1	1	0	0	2	1	1	1
H. dubium	1	1	2	0	2	1	0	1
H. erythraeum	1	1	1	0	4	2	0	1
H. furfuraceum	1	1	1	0	2&4	2	0	1
H. glaberrimum	0	1	0	0	4	1	0	1
H. griffithianum	1	1	0	0	2	1	0	1
H. laristanicum	0	1	0	0	2	1	0	?
H. latifolium	0	1	0	1	2	2	0	2
H. linifolium	1	1	2	1	4	1	0	1
H. lissonotum	0	1	2	0	2	2	0	1
H. myrtifolium	1	1	2	1	2	1	0	1
H. obtusifolium	1	1	2	0	2	2	0	1
H. patavinum	1	1	0	1	2	1	0	1
H. pilosum	1	1	2	1	5	2	0	1
H. popovii	0	1	0	0	2	1	0	1
H. ramosissimum	1	1	2	0	2	2	0	1
H. rechingeri	1	1	0	0	2	1	0	?
H. robustum	1	1	2	0	2	1	0	1
H. stapfianum	1	1	0	0	2	1	0	1
H. suaveolens	1	1	2	1	4	1	0	1
H. telephioides	1	1	2	1	4	1	0	1
H. tenue	1	1	0	0	2	1	0	1
H. thesioides	1	1	0	1	4	2	0	1
H. tuberculatum	0	1	2	1	2	1	1	1
H. versicolor	1	1	2	1	2	1	1	1
H. villosum	1	1	2	1	2	1	0	1
H. virgatum	0	1	0	0	2	1	0	1
H. viridulum	1	1	2	0	2	1	1	1
Cneoridium dumosum	0	2	0	0	2	1	0	2

Concluding remarks and future prospects

The aim of this thesis was to investigate patterns of biotic assembly in the Mediterranean region, with special emphasis on the effect of earth processes on species origins and distribution, the origin of island endemics, and the biogeographic links between the Mediterranean and Irano-Turanian regions. To address these issues, two genera of Rutaceae that met several requirements were selected as model systems: *Ruta* and *Haplophyllum*. We first carried out phylogenetic analyses for *Ruta* and closely related taxa, essential to provide a robust framework for sound biogeographic analyses (cf. **Chapter I**). The phylogenetic analyses supported the most recent circumscription of *Ruta* and allowed us to explore congruence between different classes of characters with respect to the inferred molecular phylogeny. Subsequently, by adopting an integrative approach to historical biogeography, drawing from phylogenetics, molecular dating, and ancestral range reconstruction methods that incorporate palaeo-geographic models, we proposed biogeographic scenarios for the origin and diversification of *Ruta*, and in particular its island endemics (cf. **Chapter II**). We hope that these scenarios will be compared with those of other organisms in order to identify major patterns of biotic assembly in the Mediterranean region. For example, our study demonstrated that *Ruta* invaded the Mediterranean region from the north prior to the onset of the current Mediterranean climate. Such a finding stresses the importance of the filtering of floristic elements from the so-called "Tertiary geofloras", which covered most of the Northern Hemisphere before the development of such climate (Thompson, 2005). Moreover, by projecting ancestral areas of distribution onto palaeo-geographic maps of the Mediterranean region, we were able to identify specific geologic features/processes underlying the distribution of *Ruta*. For example, our study showed that the Corso-Sardinian endemic lineage of *Ruta* entered Corsica and Sardinia via a narrow land bridge connecting these two islands with Apulia and Eurasia during the Miocene. The function of such land corridor as a route of migration between the two islands and the continent should be tested with other organisms. Finally, to explore the biogeographic links between the Mediterranean and Irano-Turanian regions, we carried out phylogenetic analyses of *Haplophyllum* (cf. **Chapter III**), which suggested that the Mediterranean representatives of the genus arrived from the east multiple times, corroborating the hypothesis that the latter region served as a key source for the colonization of the former one. This study will lay the foundations for more detailed biogeographic analyses on *Haplophyllum*, which will involve molecular dating and ancestral area reconstruction, aimed at answering the following points: i) When and where did the genus originate? ii) When did the Mediterranean representatives of

the genus originate and along which route did they colonize the Mediterranean region? For example, two main dispersal routes for the colonization of the western Mediterranean region, where a few species of *Haplophyllum* occur, from the east have been proposed: one across central Europe and one across North Africa (Sanmartín, 2003). Moreover, we hope that our work on *Haplophyllum* will be beneficial for future studies aimed at protecting the genus. Many of its species, in fact, exhibit a narrow geographic range (i.e., "narrow endemics"), a feature that makes them particularly vulnerable to extinction. Nine species of *Haplophyllum* have already been included in the Red Data Book of Iran (Jalili and Jamzad, 1999). The phylogenetic analyses of *Haplophyllum* identified a few sister pairs including a widespread species and a narrow endemic. Carrying out multiple, independent comparisons within such sister pairs will be crucial to determine the biological factors (e.g., life-history traits, dispersal mode, genetic background) that limit the geographic range of the narrow endemics as compared to the widespread species.

In recent years, as a reaction to the growing separation between historical and ecological biogeography, a new line of investigation in biogeographic research consisting in combining evidence from both disciplines has emerged (Wiens and Donoghue, 2004). Historical biogeography and ecology, in fact, are inextricably linked: the large-scale, long-term patterns of distribution studied by historical biogeographers are the product of short-term, local-scale processes (e.g., competition, dispersal, adaptation) that are generally the focus of ecologists. Conversely, patterns analysed by the latter group (e.g., community assembly, species richness) involve processes typically examined by the former one (e.g., speciation, large-scale dispersal; Stephens and Wiens, 2009). This novel approach to biogeography allows to link short-term, local processes to global processes that occur over deep evolutionary time scales (Cavender-Bares *et al*., 2009).

The advent of ecological niche modelling (ENM) and its recent integration in phylogenetic analysis (phylogenetic niche modelling, PNM) provide the necessary tools to bring historical and ecological biogeography together (e.g., Graham *et al*., 2004). Given information on the climatic conditions where a species occurs and on its environmental tolerances, an ecological niche model can be constructed and projected onto a climatic surface of an area to predict where the species is found. Assuming that the species' ecological tolerances have not changed much over time and given some information about past climate in the area of interest, the distribution of potential habitats for the focal species can be projected back in time (e.g., Vieites *et al*., 2009). Through such exercise dispersal pathways between areas that are no longer connected by suitable habitat can be predicted and areas that lacked environmental conditions favourable to the species in question (suggesting local

extinction) can be revealed (Wiens and Donoghue, 2004). For example, it still remains to be seen whether the environment present in the land corridor that connected Corsica and Sardinia with Apulia and Eurasia during the Miocene was suitable for the ancestor of the Corso-Sardinian endemic lineage of *Ruta*. Hence, ENM can illuminate the results of ancestral range reconstruction analyses by identifying past, potential, suitable habitats that match the current ecological tolerances of organisms (e.g., Smith and Donoghue, 2010).

The recent incorporation of ENM into biogeographic analysis has stirred the debate about the conditions under which ecological niches change through time. Some studies have demonstrated that species have maintained their ecological niches through time (i.e., "niche conservatism"; Peterson *et al*., 1999; Prinzing *et al*., 2001; Wiens and Graham, 2005), whereas others have shown that species have expanded their niche breadth enabling the invasion of new habitats and climatic regimes that were previously unavailable to them (i.e., "niche evolution"; Evans *et al*., 2009; Simon *et al*., 2009). As elegantly argued by Donoghue (2008), in order to assess the relative contribution of niche conservatism versus niche evolution to the assembly of the Earth's biota, one has to consider the interplay between the historical circumstances that have shaped the movement of organisms and the ease or difficulty that these have had to acquire suitable features enabling them to exploit the environment. For example, following the emergence of novel environmental conditions in an area, species will tend to be filtered into this area from neighbouring regions, where adaptations to the new environment have already evolved, only if corridors for movement are available or barriers against movement are lacking (i.e., niche conservatism; e.g., Antonelli *et al*., 2009). However, if movement across regions is restrained, the area in question will only be filled with native species that, in order to persist *in situ* under the novel environmental regime, need to expand their niche breadth (i.e., niche evolution; e.g., Simon *et al*., 2009).

It would be extremely interesting to explore the above-mentioned ideas with *Ruta*. The biogeographic analyses showed that this genus invaded the Mediterranean region from the north prior to the onset of the current Mediterranean climate. This suggests that the ancestor of *Ruta* had already evolved the characters necessary to persist under the changing Mediterranean climate before invading the Mediterranean region. To test this hypothesis, an analysis of the functional traits (i.e., those directly responsible for the growth and metabolism of a plant) of *Ruta*, within a phylogenetic framework, is of paramount importance (e.g., Ackerly, 2003, 2009; Verdú *et al*., 2003; Kraft *et al*., 2008; Simon *et al*., 2009). For example, marked differences in the composition of secondary chemical compounds, especially in furanocoumarins, have been detected among the species of *Ruta* (Milesi *et al*., 2001; Thompson, 2005; Blondel *et al*., 2010). The functional significance of secondary compounds

is well documented (e.g., Geber and Griffen, 2003) and these may play an important ecological role in *Ruta* and may have contributed to its diversification (Thompson, 2005). In order to understand what triggered the invasion of the Mediterranean region by *Ruta* and whether shifts in composition of secondary metabolites played a role, the biogeographic analyses presented in **Chapter II** should be integrated with ancestral character state reconstruction analyses of secondary chemistry to determine if such shifts have coincided with the southward range expansion of *Ruta*.

Finally, increasing concerns over the effects of global climate change on species distributions are bringing ENM methods to the forefront in order to investigate evolutionary responses to climatic shifts (e.g., Thomas *et al*., 2004; Thuiller *et al*., 2005). Species may either respond to changing environmental conditions by means of range expansions/contractions/shifts or they may keep their current range (Sommer *et al*., 2010). By projecting a species' niche model into future climatic scenarios, possible pathways of migration, areas lacking suitable climate and populations likely to undergo extinction can be predicted (e.g., Yesson and Culham, 2007). Following similar lines of reasoning as above, in the face of climate change, if suitable corridors for the dispersal of species with adequate traits are available, lineages are expected to track the habitat to which they are adapted (Donoghue, 2008). Due to the extinction risk that several species of *Haplophyllum* face, it would be important to carry out ENM analyses on this genus, especially by focusing on the climatic tolerances of the species exhibiting a narrow distribution.

Literature cited

Ackerly, D. D., 2003. Community assembly, niche conservatism, and adaptive evolution in changing environments. *Int. J. Plant Sci.* 164: S165-S184.

Ackerly, D. D., 2009. Some comments on the age, origin and evolution of the California and Mediterranean floras. *J. Biogeogr.* 36: 1221-1233.

Antonelli, A., Nylander, J. A. A., Persson, C., and Sanmartín, I., 2009. Tracing the impact of the Andean uplift on Neotropical plant evolution. *PNAS* 106: 9749-9754.

Blondel, J., Aronson, J., Bodiou, J.-Y., and Boeuf, G., 2010. The Mediterranean region: biological diversity in space and time; 2nd edition. Oxford University Press, Oxford.

Cavender-Bares, J., Kozak, K. H., Fine, P. V. A., and Kembel, S. W., 2009. The merging of community ecology and phylogenetic biology. *Ecology Letters* 12: 693-715.

Donoghue, M. J., 2008. A phylogenetic perspective on the distribution of plant diversity. *PNAS* 105: 11549-11555.

Evans, M. E. K., Smith, S. A., Flynn, R. S., and Donoghue, M. J., 2009. Climate, niche evolution, and diversification of the "Bird-Cage" evening primroses (Oenothera, sections Anogra and Kleinia). *American Naturalist* 173: 225-240.

Geber, M. A., and Griffen, L. R., 2003. Inheritance and natural selection on functional traits. *Int. J. Plant Sci.* 164: S21-S42.

Graham, C. H., Ron, S. R., Santos, J. C., Schneider, C. J., and Moritz, C., 2004. Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution* 58: 1781-1793.

Jalili, A., and Jamzad, Z., 1999. Red Data Book of Iran: a preliminary survey of endemic, rare and endangered plant species in Iran. Research Institute of Forests and Rangelands (RIFR), Tehran.

Kraft, N. J. B., Valencia, R., and Ackerly, D. D., 2008. Functional traits and niche-based tree community assembly in an Amazonian forest. *Science* 322: 580-582.

Milesi, S., Massot, B., Gontier, E., Bourgaud, F., and Guckert, A., 2001. *Ruta graveolens* L.: a promising species for the production of furanocoumarins. Plant Science 161: 189-199.

Peterson, A. T., Soberón, J., and Sánchez-Cordero, V., 1999. Conservatism of ecological niches in evolutionary time. *Science* 285: 1265-1267.

Prinzing, A., Durka, W., Klotz, S., and Brandl, R., 2001. The niche of higher plants: evidence for phylogenetic conservatism. *Proc. R. Soc. Lond. B* 268: 2383-2389.

Sanmartín, I., 2003. Dispersal vs. vicariance in the Mediterranean: historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). *Journal of Biogeography* 30: 1883-1897.

Simon, M. F., Grether, R., de Queiroz, L. P., Skema, C., Pennington, R. T., and Hughes, C. E., 2009. Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *PNAS* 106: 20359-20364.

Smith, S. A., and Donoghue, M. J., 2010. Combining historical biogeography with niche modeling in the *Caprifolium* clade of *Lonicera* (Caprifoliaceae, Dipsacales). *Systematic Biology* 59: 322-341.

Sommer, J. H., Kreft, H., Kier, G., Jetz, W., Mutke, J., and Barthlott, W., 2010. Projected impacts of climate change on regional capacities for global plant species richness. *Proc. R. Soc. Lond. B* 277: 2271-2280.

Stephens, P. R., and Wiens, J. J., 2009. Bridging the gap between community ecology and historical biogeography: niche conservatism and community structure in emydid turtles. *Molecular Ecology* 18: 4664-4679.

Thomas, C. D., *et al*., 2004. Extinction risk from climate change. *Nature* 427: 145-148.

Thompson, J. D., 2005. Plant evolution in the Mediterranean. Oxford University Press Inc., New York.

Thuiller, W., Lavorel, S., Araújo, M. B., Sykes, M. T., and Prentice, I. C., 2005. Climate change threats to plant diversity in Europe. *PNAS* 102: 8245-8250.

Verdú, M., Dávila, P., García-Fayos, P., Flores-Hernández, N., and Valiente-Banuet, A., 2003. "Convergent" traits of mediterranean woody plants belong to pre-mediterranean lineages. *Biological Journal of the Linnaean Society* 78: 415-427.

Vieites, D. R., Nieto-Román, S., and Wake, D. B., 2009. Reconstruction of the climate envelopes of salamanders and their evolution through time. *PNAS* 106: 19715-19722.

Wiens, J. J., and Donoghue, M. J., 2004. Historical biogeography, ecology and species richness. *TRENDS in Ecology and Evolution* 19: 639-644.

Wiens, J. J., and Graham, C. H., 2005. Niche conservatism: integrating evolution, ecology, and conservation biology. *Annu. Rev. Ecol. Evol. Syst.* 36: 519-539.

Yesson, C., and Culham, A., 2006. A phyloclimatic study of *Cyclamen*. *BMC Evolutionary Biology* 6: doi:10.1186/1471-2148-6-72.

Curriculum Vitae

Name:	Gabriele Salvo
Year of birth:	1981
Place of birth:	Palermo, Italy
Nationality:	Italian

EDUCATION (POST HIGH SCHOOL) ----------

2005 - 2010	Doctor of Philosophy (Ph.D.) in Natural Sciences, University of Zürich, Switzerland. Supervisor: Prof. Elena Conti.
2003 - 2004	Master of Science (M.Sc.) in Advanced Methods in Taxonomy and Biodiversity, Natural History Museum and Imperial College, London, United Kingdom. M.Sc. thesis: Phylogeny and biogeography of the African *Ficus* section *Galoglychia*. Jodrell Laboratory, Royal Botanic Gardens, Kew, United Kingdom. Supervisors: Dr. Nina Rønsted and Dr. Vincent Savolainen.
2000 - 2003	Bachelor of Science (B.Sc.) in Marine Biology and Coastal Ecology, University of Plymouth, Plymouth, United Kingdom. B.Sc. thesis: The effects of predation on the burrowing rate and feeding efficiency of the common cockle, *Cerastoderma edule*. Plymouth Marine Laboratory, Plymouth, United Kingdom. Collaborator: Dr. Chiara Romano. Supervisors: Dr. Gianluca Sará, Dr. John Bishop, and Prof. John Widdows.

INTERNATIONAL CONFERENCES----------

22 - 24 Jun 2009	Biodiversity hotspots in the Mediterranean area: species, communities, and landscape level. Cagliari, Italy. **Invited talk**: Molecular dating and biogeography of *Ruta* L. (Rutaceae): a case study from the Mediterranean basin.
8 - 12 Jan 2009	Fourth Biennial Meeting of the International Biogeography Society. Mérida, Mexico. **Contributed talk**: Molecular dating and biogeography of *Ruta* L. (Rutaceae): a case study from the Mediterranean basin.
12 - 16 May 2008	Biodiversity Research – Safeguarding the Future. Pre-"COP9" meeting. Bonn, Germany.
8 - 10 Apr 2008	GEO Biodiversity Observation Network meeting. Potsdam, Germany.
20 - 25 Aug 2007	Eleventh congress of the European Society for Evolutionary Biology. Uppsala, Sweden. **Poster**: Phylogeny of *Ruta* (Rutaceae): implications for Systematics and Biogeography.
28 - 30 Mar 2007	Speciation and Ecology 2007, Annual Symposium of the British Ecological Society. Sheffield, United Kingdom.
9 - 13 Jan 2007	Third Biennial Meeting of the International Biogeography Society. Puerto de la Cruz, Tenerife, Spain. **Poster**: Phylogeny of *Ruta* (Rutaceae): implications for Systematics and Biogeography.
25 Sep 2006	Dating the Tree of Life. University of Geneva, Geneva, Switzerland.
29 Jul - 3 Aug 2006	BSA 2006, meeting of the Botanical Society of America. Chico, California, USA. **Contributed talk**: Investigating the monophyly and biogeographic history of *Ruta* (Rutaceae) with chloroplast DNA sequence data.

31 Mar 2006	Feeding the needs of tomorrow. Plant Sciences: from basics to application. Zürich – Basel Plant Science Center. Zürich, Switzerland.
17 - 23 Jul 2005	Seventeenth International Botanical Congress. Vienna, Austria.
17 Sep 2004	London Evolutionary Research Network (L.E.R.N.) 2nd annual conference, Institute of Zoology, Regents Park, London, United Kingdom. **Contributed talk**: Phylogeny and Biogeography of the African *Ficus* section Galoglychia.

OTHER TALKS----------

11 Dec 2009	Polo Botanico "Hanbury", Dip. Te. Ris., University of Genova, Italy. **Invited talk**: L'origine geografica e temporale degli endemismi insulari nel Mediterraneo.

TEACHING EXPERIENCE----------

Teaching assistant in the following courses at the University of Zürich:
Molecular Evolution and Phylogenetics (BIO 222). Winter semester 2006 and 2008.
Biodiversity: Systematics, Paleontology, Evolution (BIO 113). Winter semester 2008 and 2009.

SUPERVISION----------

Sara Manafzadeh – Tarbiat Moallem University, Tehran, Iran. Project: Systematics, evolution, and conservation of the Irano-Turanian genus *Haplophyllum* (Rutaceae; flowering plants). University of Zürich, Oct 2007 - Oct 2009.

FIELD WORK----------

May - Jun 2010	Turkey and Georgia. Collecting species of *Haplophyllum* (Rutaceae) and *Primula* (Primulaceae).
May 2009	Turkey, Syria, and Jordan. Collecting species of *Haplophyllum* (Rutaceae).
May - Jun 2008	Iran and Uzbekistan. Collecting species of *Haplophyllum* (Rutaceae).
May - Jun 2006	Greece and Turkey. Collecting species of Rutaceae, Araceae, and Boraginaceae.
May 2005	Corsica (France) and Sardinia (Italy). Collecting species of Rutaceae, Araceae, and Boraginaceae.
Nov 2001	Nova Scotia, Canada. Coastal Ecology field trip.

PUBLICATIONS----------

Salvo, G., Manafzadeh, S., Ghahremaninejad, F., Tojibaev, K., Zeltner, L., and Conti, E. Phylogeny, morphology, and biogeography of *Haplophyllum* (Rutaceae), a species-rich genus of the Irano-Turanian floristic region. Accepted for publication in *Taxon*.

Salvo, G., Ho, S.Y.W., Rosenbaum, G., Ree, R., and Conti, E. Tracing the temporal and spatial origins of island endemics in the Mediterranean region: a case study from the citrus family (*Ruta* L., Rutaceae). *Systematic Biology* doi: 10.1093/sysbio/syq046.

Salvo, G., Bacchetta, G., Ghahremaninejad, F., and Conti, E., 2008. Phylogenetic relationships of Ruteae (Rutaceae): new evidence from the chloroplast genome and comparisons with non-molecular data. *Molecular Phylogenetics and Evolution* 49: 736-748.

Rønsted, N., **Salvo, G.**, and Savolainen, V., 2007. Biogeographical and phylogenetic origins of African fig species (*Ficus* section *Galoglychia*). *Molecular Phylogenetics and Evolution* 43: 190-201.

OTHER ACTIVITIES----------

Member of the organization committee of the international scientific conference: Origin and evolution of biota in Mediterranean climate zones: an integrative vision, University of Zürich, 14 - 16 July 2007.